# OF POLES AND ZEROS

# MODERN APPROACHES IN GEOPHYSICS

VOLUME 15



*The titles published in this series are listed at the end of this volume.*

# OF POLES AND ZEROS

## Fundamentals of Digital Seismology

*2nd Edition*

*by*

FRANK SCHERBAUM

*Institut für Geowissenschaften,*
*Universität Potsdam,*
*Germany*

Including a Java Applet of the
**Digital Seismology Tutor**
by Elke Schmidtke and Frank Scherbaum

SPRINGER SCIENCE+BUSINESS MEDIA, B.V

A C.I.P. Catalogue record for this book is availbale from the Library of Congress.

**Additional material to this book can be downloaded from http://extras.springer.com**

ISBN 978-0-7923-6835-9 ISBN 978-1-4020-6861-4 (eBook)
DOI 10.1007/978-1-4020-6861-4

---

*Printed on acid-free paper*

First published 1996
Reprinted 2000
Second edition 2001

# Contents


| | | |
|---|---|---|
| | **Preface to the second edition** | **VII** |
| | **Preface to the first edition** | **IX** |
| | **Foreword by Eric Bergman** | **XI** |
| **1** | **Introduction** | **1** |
| **2** | **RC Filter** | **12** |
| 2.1 | The system diagram | 12 |
| 2.2 | The differential equation | 13 |
| 2.3 | The frequency response function and the Fourier transform | 14 |
| 2.4 | The transfer function and the Laplace transform | 21 |
| 2.5 | The impulse response function | 22 |
| 2.5.1 | The condition for stability | 29 |
| 2.6 | The step response function | 29 |
| 2.7 | The frequency response function and the pole position | 31 |
| 2.8 | The difference equation | 36 |
| 2.9 | Review RC filter | 37 |
| **3** | **General linear time invariant systems** | **39** |
| 3.1 | Generalization of concepts | 39 |
| 3.2 | Graphical estimation of the frequency response function | 42 |
| 3.3 | The phase properties of general LTI system | 44 |
| 3.4 | The interpretation of the frequency response function | 45 |
| **4** | **The seismometer** | **48** |
| 4.1 | The solution for simple initial conditions | 51 |
| 4.1.1 | Underdamped case | 54 |
| 4.1.2 | Overdamped case | 55 |
| 4.1.3 | Critically damped case | 56 |
| 4.1.4 | Comparison | 56 |
| 4.2 | The determination of the damping constant | 57 |
| 4.2.5 | The electromagnetic sensor | 58 |
| 4.3 | The frequency response function | 60 |
| 4.3.6 | The transfer function | 63 |
| **5** | **The sampling process** | **66** |
| 5.1 | The sampling of analog data | 66 |

| | | |
|---|---|---|
| **6** | **Analog-to-digital conversion** | **73** |
| 6.1 | The principle of analog to digital conversion | 73 |
| 6.2 | Increasing dynamic range and resolution | 81 |
| 6.2.1 | Gain ranging | 81 |
| 6.2.2 | Oversampling | 84 |
| **7** | **From infinitely continuous to finite discrete** | **95** |
| 7.1 | Fourier series for periodic continuous-time signals (CTFS) | 95 |
| 7.2 | Fourier transform for aperiodic continuous-time signals (CTFT) | 97 |
| 7.3 | Fourier transform of discrete-time signals | 99 |
| 7.4 | Fourier series for periodic discrete-time signals | 103 |
| 7.5 | The Discrete Fourier Transform (DFT) | 104 |
| 7.6 | The z-transform and the discrete transfer function | 107 |
| 7.7 | The inverse z-transform | 112 |
| 7.8 | The z-transform and the Discrete Fourier Transform (DFT) | 113 |
| 7.9 | The convolution of sequences | 114 |
| **8** | **The digital anti-alias filter** | **117** |
| 8.1 | Removing the maximum phase component of a FIR filter | 122 |
| 8.1.1 | The difference equation | 124 |
| 8.1.2 | Practical considerations | 125 |
| **9** | **Inverse and simulation filtering of digital seismograms** | **137** |
| 9.1 | Stability problems | 139 |
| 9.2 | Implementation | 150 |
| 9.2.1 | From the spectral to the time domain | 151 |
| 9.2.2 | FIR simulation filters | 153 |
| 9.2.3 | Bilinear transform | 156 |
| 9.2.4 | Least squares approach | 163 |
| 9.3 | Review inverse and simulation filtering | 164 |
| **10** | **The measurement of wavelet parameters from digital seismograms** | **166** |
| 10.1 | The determination of ground motion amplitudes | 168 |
| 10.1.1 | Review amplitude determination | 182 |
| 10.2 | The properties of seismic onsets | 183 |
| 10.2.1 | Onset order | 184 |
| 10.2.2 | Onset polarity | 188 |
| 10.2.3 | Filter delays and onset time | 188 |
| 10.3 | Rise time and pulse duration | 195 |
| 10.4 | Signal moment | 201 |
| | **References** | **204** |
| | **Appendix: Solution to problems** | **208** |
| | **Index** | **265** |

***CD-ROM including Java Applet Digital Seismology Tutor: inside back cover***

# Preface to the second edition

Since the publication of the first edition in 1996, I have received a lot of feedback from students and colleagues which has motivated me to rethink the way this book can be used in courses or for self-study. I believe it to be essential that the examples and homework problems can be reproduced and modified in an easy way. For this purpose, the second edition contains a CD with the *Digital Seismology Tutor,* a Java applet (including the source code) written by Elke Schmidtke and myself. The *Digital Seismology Tutor* (DST) is maintained and updated by Elke Schmidtke (University of Potsdam, email: eschmidt@geo.uni-potsdam.de; http://www.geo.uni-potsdam.de). It contains all the relevant educational tools from PITSA (Scherbaum and Johnson, 1993), which was used to create the examples in the first edition. Consequently, the *Appendix: Solutions to problems* is now completely based on the *Digital Seismology Tutor.* Hence the examples and problems from this book can now be reproduced and/or modified on any platform capable of running Java.

In addition to including the *Digital Seismology Tutor*, small errors and typos which have been caught in the first edition, have been removed. In this context, I am indebted to Angela Sachse for her efforts. Regarding the contents, chapter 6 has been modified considerably to cover the topic of sigma-delta modulation including new problems/exercises and solutions.

Potsdam, October 2000
Frank Scherbaum

# Preface to the first edition

This book is aimed primarily at observational seismologists dealing with the acquisition and analysis of digital seismic data. Seismic waveforms and the parameters commonly extracted from them such as onset times, amplitudes, and pulse durations, etc., are strongly influenced by the effects of numerous filters both within the earth and within the recording system. As seismologists, we must be aware of these effects and the constraints they impose on the results of our analysis.

In addition, with the advent of numerous software tools for the processing of digital seismograms, observational seismologists have unprecedented power in extracting information from seismic records. These tools are often based on sophisticated theoretical aspects of digital signal processing, which to be used properly, must be understood.

Signal processing is a wide field and as seismologists interested in this topic, we are confronted with literally hundreds of books from which we have to extract bits and pieces of information. This book is not meant to compete with any of them. On the contrary, a number of them - which are given in the references - have been quite influential in shaping this text. The primary goal of this course is to concentrate on those aspects of digital signal processing that are relevant to the interpretation of seismograms. Therefore it focuses on the building blocks of modern seismic recording systems and step by step develops the concepts needed for the quantitative understanding of their properties and their effects on seismic waveforms.

The text is accompanied by numerous examples, homework problems and solutions, which all are an essential part of this course. They are not only aimed at strengthening the theoretical understanding, but also provide a practical feeling for some of the relevant phenomena. Most of them have been produced using PITSA, a program originally written by myself and Jim Johnson (Scherbaum and Johnson, 1993). A PC version of the program is available as Vol. 5 of the IASPEI Software Library through the Seismological Society of America, 201 Plaza Professional Building, El Cerrito, CA 94530, USA (Phone: 510-525-5474: Fax: 510-525-7204). A UNIX version for Sun workstations can be obtained through anonymous ftp from IRIS (dmc.iris.washington.edu). Although desirable, access to the PITSA program is not absolutely necessary to follow the material presented. Readers with access to programs such as MathCAD[1], Mathematica[2], or

Maple[1] will find that they can reproduce all of the essential examples with only moderate programming efforts. For readers who completely lack access to any of these programs, PITSA screendumps for the most important steps of the solutions are given in "" starting on page 208. PITSA was primarily designed for research purposes and the application in day-to-day station-based seismological analysis. With the analysis tools becoming theoretically more and more demanding, however, we began to add some educational features, mainly to test the behaviour of certain algorithms and to teach ourselves the proper application. This led to the first version of a *Short Course on the First Principles of Digital Signal Processing for Seismologists* (Scherbaum, 1993) followed by *Basic Concepts in Digital Signal Processing for Seismologists* (Scherbaum, 1994).

Since 1991, this course has been taught to different audiences in a number of different countries, most commonly as training workshop for the upcoming international Seismological Observing Period (ISOP) project (Dornboos et al., 1990; see also the foreword by Eric Bergman). I have benefited enormously from the feedback and the enthusiasm of the participants of these courses. Based on the experience from these workshops, I have started in 1994 to extend the material to include topics such as "Inverse and Simulation filtering of digital seismograms" and "The measurement of wavelet parameters from digital seismograms". I have also revised the original chapters to make them more self-contained theoretically.

I am indebted to the students at the University of Munich who involuntarily had to act as test subjects (and still do) for the teaching concept behind this course. I have benefited from numerous discussions on the topics of this book with Axel Plesinger, Miroslav Zmeskal, Jim Johnson, Joachim (Jo) Wassermann, Dieter Stoll and Erhard Wielandt. Wang-Ping Chen, Eric Bergman, Axel Plesinger, Miroslav Zmeskal, Nicolas Deichmann and Peggy Hellweg contributed considerably to the shape of this book by critical reading and offering helpful comments. I owe special thanks to Peggy Hellweg for her attempts to improve the clarity of the text.

The cover illustration shows the FIR filter of the SIL data acquisition system (Stefánsson et al., 1993).

Munich, January 1996
Frank Scherbaum

---

1. MathCAD is a registered trademark of MathSoft, Inc.
2. Mathematica is a registered trademark of Wolfram Research, Inc.
1. Maple is a registered trademark of Waterloo Maple Software.

# Foreword

by
Eric Bergman
(ISOP coordinator)

The idea for the International Seismological Observing Period (ISOP) project was born in the mid-1980s as a mechanism to focus attention on the needs of the infrastructure that supports global seismology: the hundreds of seismic stations and observatories, the protocols and agreements that govern data collection, processing, analysis, distribution, and archiving, and especially the thousands of people who make this basically ad hoc system work. The ISOP concept received official endorsement at the IASPEI General Assembly in Istanbul in 1989, and achieved some sort of reality in September, 1990, when I came to the U.S. Geological Survey's National Earthquake Information Center in Golden, Colorado, to open the ISOP Office, in collaboration with Bob Engdahl. While we had a number of grand ideas about the project, at that time none of them had really coalesced into a practical program of activities. This situation began to change one day in December, 1990, when Frank Scherbaum came to visit me in Golden. I had only been in residence for a few months as ISOP Coordinator, and Frank was the first person ever to come to Golden specifically to discuss ISOP.

Bob Engdahl had met Frank at the IASPEI Regional Assembly in Kenya a few months earlier and thought he had potential as a good "ISOPer". Fortunately, Jim Johnson, Frank's co-author on the PITSA software, lived in Boulder and it was easy to arrange a meeting while he was in town to work with Jim. At that time he was making his first attempts to publicize PITSA to the rest of the seismological community. In addition to discussing how PITSA might be used in the ISOP project, he gave me an early draft of what was to become the first version of the present book, the booklet "First Principles of Digital Signal Processing for Seismologists". This meeting launched the concept - which has become the organizing principle of ISOP training courses - of a unified package of educational material and tutorials closely linked to interactive analysis software which was designed to support the educational function as well as research and limited production functions.

Shortly after Frank's visit, and largely in response to the ideas we discussed then, the ISOP project sponsored a major international workshop on the subject of station-based analysis of digital waveform data. This workshop, held in Denver in March, 1991, helped crystallize many of the ideas which have since come to be thought of as the core of the

ISOP project. It was a tempestuous and stimulating workshop, attended by a large group of experienced seismologists (including Frank) who brought a variety of viewpoints to the discussions. In the end we managed to bring the workshop to a satisfactory conclusion: the participants agreed on a statement of general principles regarding the importance of developing tools to support station-based analysis of waveform data, and some broad design guidelines for software were established. Equally important was acknowledgement of the vital importance of providing training and education in the use of those software tools. Although it was not the original intent, this workshop largely defined ISOP's central theme.

The 1991 Denver workshop brought another important figure, Axel Plesinger, into the ISOP story. With colleagues in Prague and Germany, Axel had for many years been developing a set of rigorous and sophisticated algorithms for processing broadband digital data, especially codes to address the simulation problem: to produce a digital record which would have been produced on a different kind of instrument. This turned out to be a very important issue for ISOP, since the goal of making coordinated, standardized measurements of secondary phases at many stations worldwide required a reliable method to simulate standard instruments, regardless of the original seismograph system.

In 1992, Axel and I obtained funding under the U.S.-Czech Science and Technology Program to develop his codes into an integrated package - PREPROC - for pre-processing raw digital waveform data, including simulation of standard instrument responses. Much of the final algorithm development and coding was done by Miroslav Zmeskal, and significant contributions also came from Mila Musil and Jan Zednik. The Czech group also wrote a manual to accompany the software, one that would go well beyond a simple user's guide and cover many of the fundamentals of digital signal processing relevant to the code. In this sense the PREPROC Manual bears a strong resemblance to Frank Scherbaum's booklet "First Principles..." and the present book. In fact considerable effort by Frank and Axel went toward making the two volumes complimentary and consistent with one another, but not redundant. As a witness to some of their "reconciliation" sessions, I can confirm that each made significant contributions to the other's manuscript.

Since our first meeting in 1990, Frank has gone on to refine both PITSA and the lecture notes which became the "First Principles...". These resources have now been used in several ISOP training courses, and the tendency for his efforts to be thought of as part of the ISOP project has grown. I am very proud of the support and opportunities that the ISOP project has given Frank to carry out his work over the last five years, but I believe he would have done it by himself in any case. Frank's efforts in ISOP have been motivated by a belief that seismologists, even those working in observatories in what is often referred to as "routine practice", should have the tools and training to carry out their duties at the highest possible level of practice not just because that will tend to provide the best data for research seismologists, but because it is best for the health of the science as a whole.

View this book, therefore, as one aspect of a range of activities and products produced by

a community of seismologists who have chosen to invest some of their time and energies in cultivation of the fundamental resources - the people - of seismology as a globally-conducted observational science. They organize or donate their time to lecture at training courses, support programs for technology transfer and visiting scientists, write and distribute useful and reliable software, and write tutorials and training materials. They take the time to learn about the real needs of seismologists in other organizations and countries, and they try to find ways to meet those needs. That the ISOP project has helped to attract, organize, and encourage seismologists of this temperament may be it's highest achievement.

Frank Scherbaum's contributions to ISOP have been larger than he would ever want to admit, not least in helping define what a "good ISOPer" is, but they are not so large that nothing remains to be done. Over the last five years more than a few like him have "walked into my office" in one way or another. I'm looking forward to the next one.

Golden, October 1995
Eric Bergman

CHAPTER 1

# Introduction

Seismologists try to obtain information about physical processes within the earth by recording and analysing seismic ground motion. Historically, their attention has been focused on the earthquake source. Since, however, the ground motion recorded at a seismic station on the earth's surface differs considerably from seismic signals generated by the earthquake source a major problem in seismology is the separation of the original source signal from other effects.

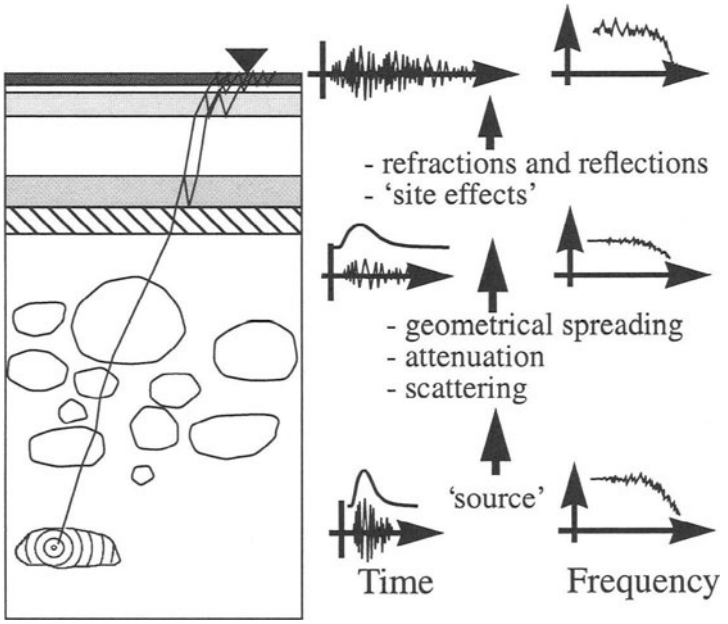



**Fig. 1.1** Signal distortion during wave propagation from the earthquake source to the surface.

In Fig. 1.1, some of the conceivable changes in the waveform along the propagation path are sketched schematically. Attenuation causes frequency dependent reduction of the amplitudes and phase shifts. Scattering will produce complicated superpositions of wavelets with different paths, and reverberation in shallow sedimentary layers will cause frequency dependent amplification. Finally, the recording system and the sampling process produce additional signal distortion.

Fig. 1.2 illustrates the influence of the recording system on the observed waveform by showing the vertical component record of the Izmit earthquake of August 17, 1999 as it would be recorded on three different instruments in a distance of 29.8°. Each instrument distorts the incoming signal in a different way, emphasizing different frequency components of the signal.

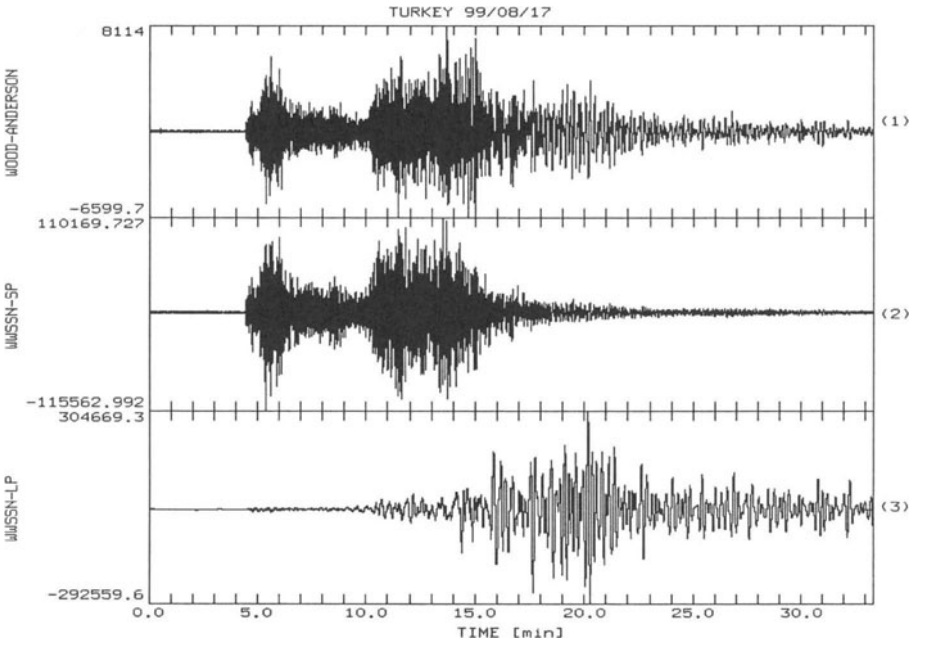



**Fig. 1.2** Vertical component record of the Izmit earthquake in Turkey (1999/08/17) recorded at station MA13 of the University of Potsdam during a field experiment in Northern Norway. Shown from top to bottom are the vertical component records for a: Wood-Anderson, a WWSSN SP, and a WWSSN LP instrument simulation.

Before we can even begin to think about the interpretation of recorded seismic signals in terms of properties of the source and/or the earth, we have to understand and possibly correct for the effects caused by the recording process. In the following chapters we will see how these effects can be understood and modelled quantitatively.

In addition to the effects of the seismometers, we have to understand the limitations imposed on the data by the sampling process. For maximum resolution, current state-of-the-art recording systems often make use of oversampling/decimation techniques. For reasons which will be discussed later, digital lowpass filters with two-sided impulse responses similar to the top trace in Fig. 1.3 are commonly used in this context. As a consequence, the onset times for very impulsive seismic signals may be obscured by precursors and become difficult, if not impossible, to determine, especially for automatic phase pickers.

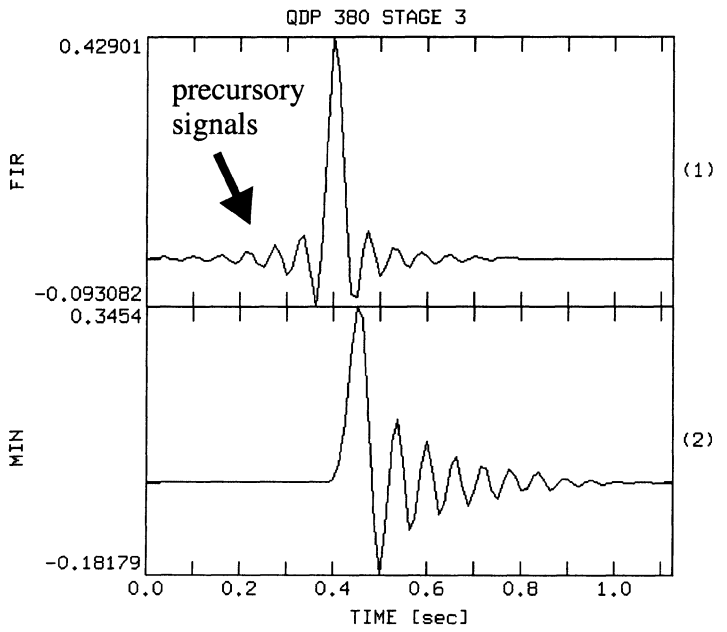



**Fig. 1.3** Impulse response of stage 3 of the two-sided decimation filter (with zero-padding) incorporated in the Quanterra QDP 380 system (top trace). The bottom trace shows a filter response with an identical amplitude but different phase response.

In chapter 8 *The digital anti-alias filter*, we will see how to treat this problem theoretically as well as practically. We will learn how to remove the acausal (left-sided) portion of such a filter response from seismic signals for the determination of onset times. In other words, we will see how to change the two-sided filter response shown in the top trace of Fig. 1.3 into a right-sided equivalent, as shown in the bottom trace of Fig. 1.3.

In the final introductory example, we will look at a typical sequence of signal processing steps to determine seismic source parameters from local earthquake records (Fig. 1.4).

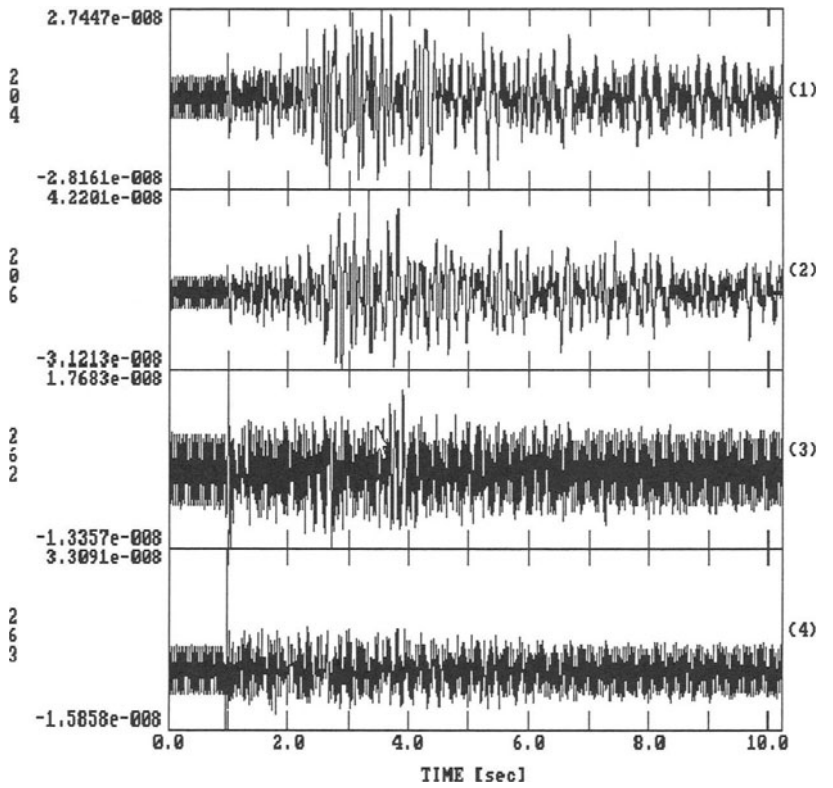



**Fig. 1.4** Vertical component, velocity records of stations 204, 206, 262, and 263, respectively of event 4 of the Chalfant Valley aftershock sequence (cf. Luzitano, 1988). Notice the monochromatic noise that has been artificially superimposed on the data to simulate the effects of crosstalk.

Fig. 1.4 shows four short-period, vertical component records for an aftershock of the 1986 Chalfant Valley earthquake recorded at temporary stations of the USGS. Monochromatic 60 Hz signals have been artificially superimposed on the original records in order to simulate the potential effect of crosstalk. After the correction for the preamplifier gains, typically the first step of digital signal processing would remove this and other kinds of unwanted 'noise' from the data. The techniques covered in this course will enable us to design and apply simple filters for this purpose. After the noise has been removed, the records from stations 204, and 262 look like Fig. 1.5.

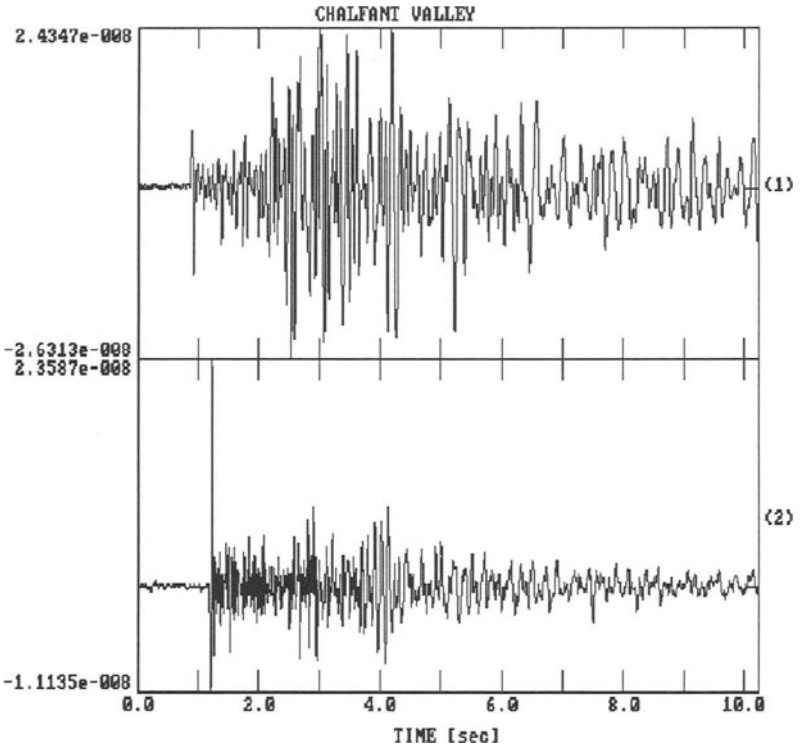



**Fig. 1.5** Velocity records of stations 204, and 262, respectively, of event 4 of the Chalfant Valley aftershock sequence after noise removal.

By applying signal processing techniques to the recorded data we would ideally like to obtain either an approximation to the true ground motion or at least a simulation of records of some standardized instruments. In technical terms, the first problem is called *signal restitution*, while the second one is known as the *instrument simulation* problem. For both of these tasks, the behaviour of the recording system must be described quantitatively. Two of the most powerful tools in this context are the concepts of the *transfer function* and the *frequency response function*. The modulus of the latter provides a convenient way to visualize the frequency dependent amplification and damping of seismic signals by a specific instrument. The theoretical frequency response function (modulus) for the recording system used to acquire the data shown above is shown in Fig. 1.6.

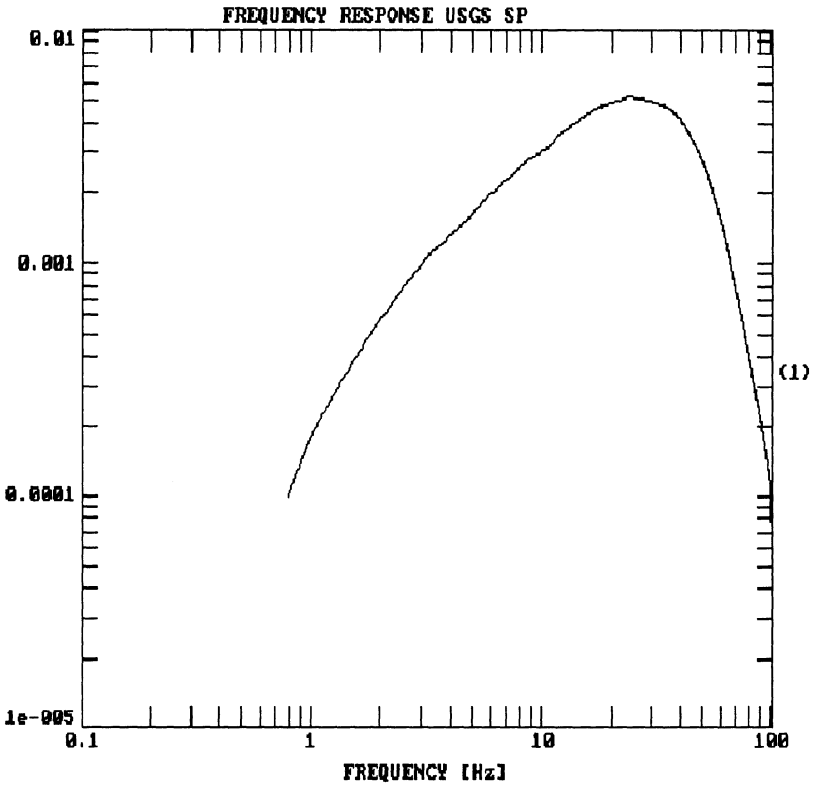



**Fig. 1.6** The theoretical frequency response function (modulus) for the USGS short-period refraction system (normalized response to ground displacement). After Luzitano (1988).

When we can describe the characteristics of seismic recording systems, we will learn how they interact with the ground motion. This will lead to *convolution*, one of the essential topics of system theory. In the context of restitution and/or simulation, however, we are even more interested in the inverse process, *deconvolution*. There are numerous techniques for convolution and deconvolution, each with its own advantages and drawbacks. These topics are discussed in chapter 9 *Inverse and simulation filtering of digital seismograms*. For the purpose of this example, we will use the technique of *spectral division* to illustrate some of the basic effects. If we apply this method to the data in Fig. 1.5, we obtain the records shown in Fig. 1.7.

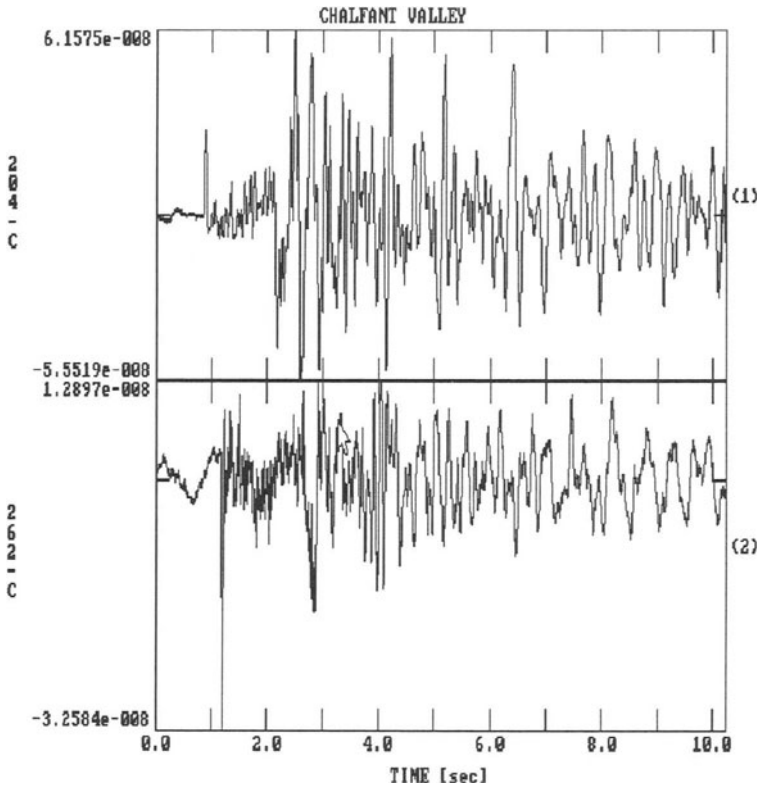



**Fig. 1.7** Instrument corrected displacement records of the signals shown in Fig. 1.5.

Once the signals have been corrected for the instrument response, they can be inverted to determine seismic source parameters. From this kind of records, we conventionally try to determine estimates of the seismic moment $M_o$, which is proportional to the product of source size (radius) and source dislocation, the stress drop, and attenuation properties along the ray path. In this context, it is quite common to do the interpretation in the *spectral domain*, by calculating the Fourier transform of the data. If we do this for the P wave portion of the top trace in Fig. 1.7, we obtain the *spectrum* shown in Fig. 1.8. Its shape is more or less typical of a displacement spectrum, with an essentially flat low frequency level (here roughly between 1 and 10 Hz), and a high frequency region in which the spectral amplitude decays rapidly with increasing frequency (here above 10 - 20 Hz). The transition region is commonly described by the *corner frequency*, $f_c$, here defined as the frequency at which the spectral amplitude of the source signal is 1/2 of the plateau value (cf. equation (1.2)). In a double logarithmic plot, the corner frequency appears roughly at the intersection of two asymptotes: one for the plateau and the other for the high-frequency region of the spectrum.

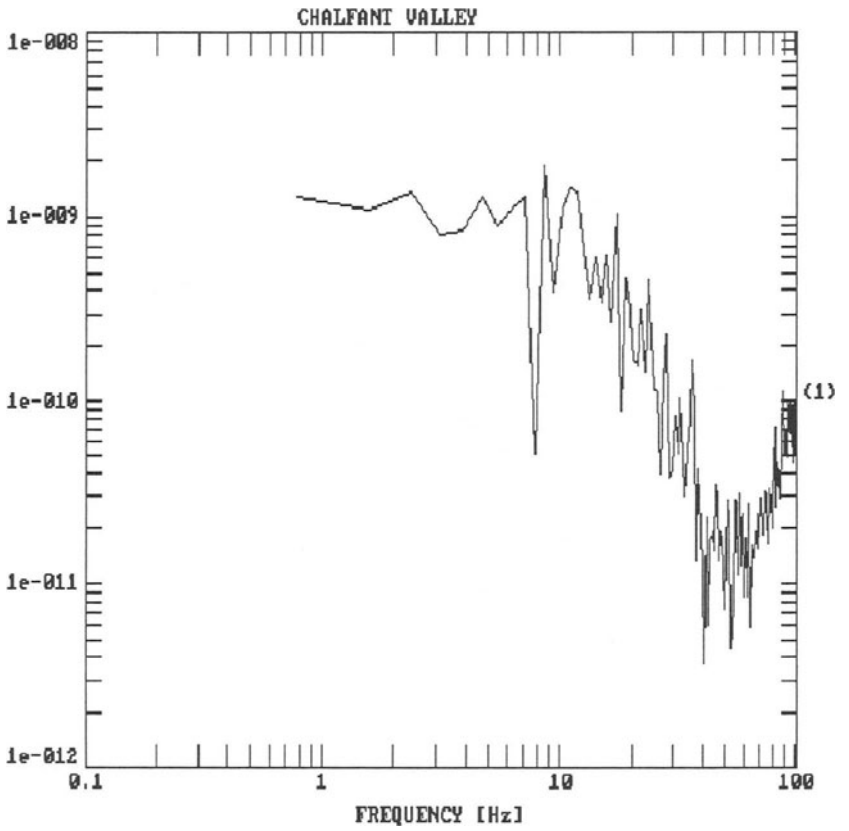



**Fig. 1.8** Displacement spectrum for the P- wave portion of the instrument corrected displacement record of station 204 (top trace in Fig. 1.7).

Under very general assumptions, the displacement spectrum of an observed seismic P or S wave signal can be modelled as the product of different spectral factors:

$$S(f) = A(f) \cdot I(f) \cdot R(f) \cdot B(f) \tag{1.1}$$

Here $S(f)$ represents the observed spectrum. On the right hand side, the contributing factors are the far field source spectrum $A(f)$, the instrumental response $I(f)$, the site response $R(f)$, and the attenuation spectrum $B(f)$. The far field displacement source spectrum in a homogeneous whole space is commonly modelled as:

$$A(f) = \frac{M_0 \cdot R(\theta, \phi)}{4 \cdot \pi \cdot \rho \cdot s \cdot v^3} \cdot \frac{f_c^\gamma}{f_c^\gamma + f^\gamma} \tag{1.2}$$

Here $M_0$ denotes the seismic moment, $R(\theta, \phi)$ the radiation pattern, $\rho$ the density of the earth at the source, $s$ the hypocentral distance, $v$ the P-, or S- wave velocity, depending on which signal is being analysed, $f_c$ the source corner frequency, $\gamma$ the high frequency decay factor (assumed to be 2 or 3 for P waves), and $f$ the frequency, respectively. The attenuation spectrum $B(f)$ is taken to be:

$$B(f) = e^{-\pi f s/(vQ)} = e^{-\pi f t/Q} \tag{1.3}$$

with $t$ being the traveltime. Once the instrument response has been removed from the observed signal equation (1.1) reduces to

$$S(f) = A(f) \cdot R(f) \cdot B(f) \tag{1.4}$$

If we know or can make reasonable assumptions about the attenuation factor $Q$, the elastic parameters of the medium, and the site spectrum $R(f)$, equation (1.1) can directly be used to invert the spectrum in Fig. 1.8 for seismic source parameters. The inverse theory necessary to perform this task is, however, beyond the scope of this text. Readers interested in the topic of inversion should consult, for example, the books by Menke (1984), Tarantola (1987) and Hjelt(1992).

In the context of seismogram interpretation, measurements are usually made on records from either real or simulated seismic instruments rather than on records of true ground motion. In general, only a small number of signal parameters are analysed. Some of these are sketched in Fig. 1.9. In chapter 10 *The measurement of wavelet parameters from digital seismograms* we will discuss how these parameters are affected by the properties of standard seismic recording systems.

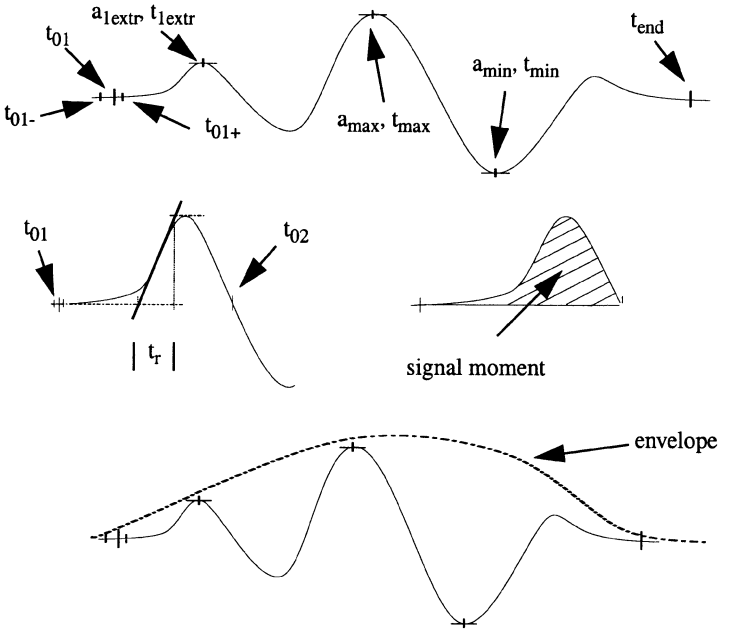



**Fig. 1.9** Commonly determined signal parameters such as onset times and amplitudes, rise time, signal moment or envelope.

The signal processing necessary to recover information about the earth from seismograms is very powerful, but also has limitations. This must be taken into account during analysis, and particularly during the interpretation of data. The concepts underlying the processing of seismic signals are the general principles of signal processing or the theory of filters and systems. We make the following general definition:

> • *Definition* — Filters or systems are, in the most general sense, devices (in the physical world) or algorithms (in the mathematical world) which act on some input signal to produce a - possibly different - output signal.

Systems are often visualized as block diagrams with a line or an arrow representing the input signal, and another representing the output signal (Fig. 1.10).

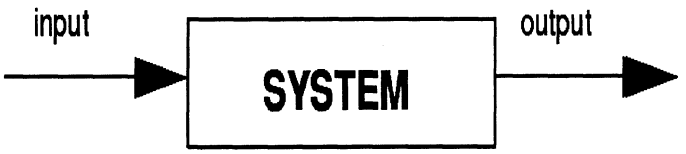



**Fig. 1.10** Block diagram of a system

We can generate a basic model of the factors affecting a seismic signal using very simple assumptions. A possible model is shown in Fig. 1.11. As seismologists, we are especially interested in isolating the left half of the diagram for interpretation in terms of geophysical models.

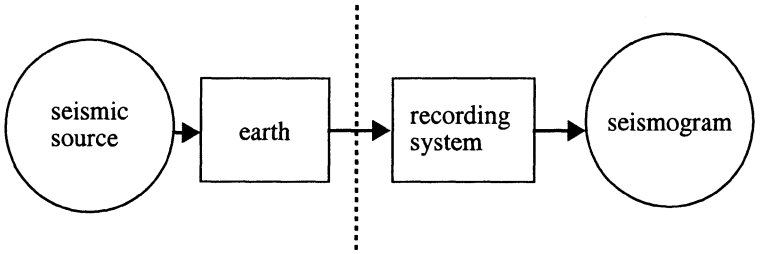



**Fig. 1.11** System diagram of the generation of a seismogram

If it is only used to display the logical structure of a sequence of physical processes, the concept of a system is of little help. To make real use of this concept, we must first learn how to quantitatively describe the properties of systems and how they act on signals. If a system meets the simple requirements of *linearity* and *time invariance*, we will see that we do not need to know what is physically going on inside of the system box. All we need to know in order to describe the properties of such a system and how it changes an arbitrary input signal is to know how it changes some very simple input signals like impulses or sinusoids. It will turn out that systems are most conveniently described in the so-called 'frequency' domain which is related to the time domain via the Fourier- or Laplace-transform.

CHAPTER 2

# RC Filter

A very simple filter will highlight some basics from the theory of systems. This filter consists of a capacitor $C$ and a resistor $R$ in series (Fig. 2.1).

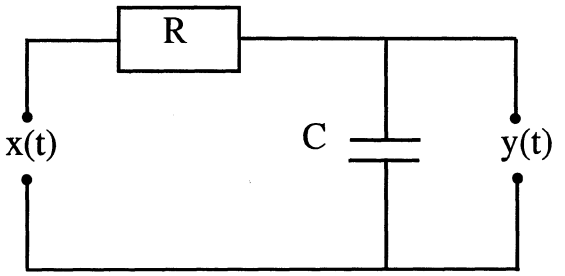



**Fig. 2.1** RC filter.

Because it is easy to understand we can use it to compare different ways of quantifying the system's physical and non-physical properties.

## 2.1 The system diagram

Let us start by making use of the knowledge we have so far about the theory of filters and make a graphical representation of the RC filter as a simple block diagram (Fig. 2.2). This tells us that for the input signal $x(t)$, we obtain an output signal which we call $y(t)$.


**Fig. 2.2** System diagram for the RC filter in Fig. 2.1.

## 2.2 The differential equation

If we apply a time-dependent input voltage $x(t)$ to the left terminals of the circuit, a current $I$(t) will flow through the resistor $R$ and across the capacitor $C$. The voltage difference measured across the resistor will be $RI(t)$. If we call the voltage at the two pins on the right $y(t)$, the overall voltage balance is:

$$RI(t) + y(t) = x(t) \tag{2.1}$$

The current is controlled by the capacitance $C$:

$$I(t) = C\dot{y}(t) \tag{2.2}$$

Inserting equation (2.2) into equation (2.1) we obtain the differential equation describing the voltage in the electric circuit:

$$RC\dot{y}(t) + y(t) - x(t) = 0 \tag{2.3}$$

This differential equation describes the physical properties of the circuit. Equation (2.3) is an example of a first order linear differential equation. For that reason, we call the corresponding system **a linear system.** Linear systems have the property that if $y_1(t)$and $y_2(t)$ are the output signals corresponding to the input signals $x_1(t)$and $x_2(t)$, respectively, the input signal

$$x_3(t) = \alpha_1 x_1(t) + \alpha_2 x_2(t) \tag{2.4}$$

will produce the output signal:

$$y_3(t) = \alpha_1 y_1(t) + \alpha_2 y_2(t) \tag{2.5}$$

We also call this system **time invariant** since the properties of the filter (as described by the resistor and capacitor) are assumed to be constant in time.

## 2.3 The frequency response function and the Fourier transform

For a zero input signal ($x(t) = 0$) we obtain the homogeneous equation

$$RC\dot{y}(t) + y(t) = 0 \tag{2.6}$$

We can easily see by restitution into equation (2.6) that

$$y(t) = -\frac{1}{RC}e^{-t/(RC)} \tag{2.7}$$

is a solution. However, we want to know how the filter acts on an arbitrary input signal. Fortunately, an arbitrary function can be described as a superposition of harmonics under very general conditions (Fourier series or Fourier integrals, see (2.10)). If we know the output signal of a linear system for an harmonic input signal, we can obtain the solution for an arbitrary signal by superposition. We will therefore briefly review a few essential properties of the Fourier transform. Comprehensive treatment of the underlying theory can be found in Papoulis (1962), Oppenheim and Schafer (1989), or Kraniauskas (1992).

The Fourier transform is defined by the following equation:

$$\mathcal{F}\{x(t)\} = X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}\, dt \tag{2.8}$$

It is common practice to write the frequency in terms of angular frequency $\omega = 2\pi f$, and since (2.8) describes a complex quantity to write instead:

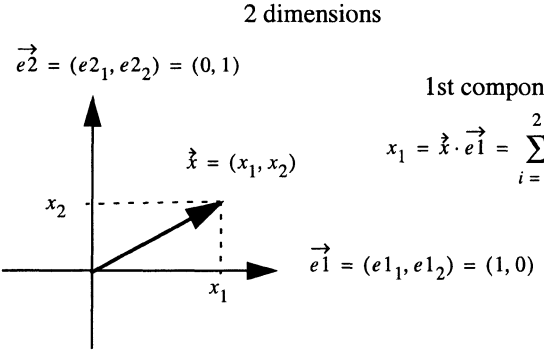
$$X(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\, \mathrm{dt} \tag{2.9}$$

The inverse transformation is:

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\omega)e^{j\omega t}\, d\omega \tag{2.10}$$

One reason for using $j\omega$ is that this will allow us to see the Fourier transform as a special case of the Laplace transform which will be introduced in the following chapter. The Fourier transform is a linear transform, which means that the transform of the scaled sum of two signals is the scaled sum of the individual transforms.

Intuitively we can think of the Fourier transform as mapping a 'time' signal $x(t)$ from the 'time domain' onto a 'frequency' signal $X(f)$ in the 'frequency domain' similar to projecting a vector onto a basis vector of a coordinate system in order to obtain a particular component (Fig. 2.3).




1st component in 3 dimensions

$$x_1 = \vec{x} \cdot \vec{e1} = \sum_{i=1}^{3} x_i \cdot e1_i$$

k-th component in n dimensions

$$x_k = \vec{x} \cdot \vec{ek} = \sum_{i=1}^{n} x_i \cdot ek_i$$

---

function space

$$i \to t \qquad k \to f \qquad n \to \infty \qquad \vec{e} \to e^{-j2\pi ft} \qquad \sum_{i=1}^{n} \to \int_{-\infty}^{\infty} dt$$

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft}\, dt \qquad \text{'frequency component'}$$

**Fig. 2.3** Fourier transform as 'projection' onto general harmonic functions.

In other words, the spectrum can be thought of representing the components of a 'time signal' in terms of harmonic functions. Of course, the meaning of 'time-domain $t$' and 'frequency domain $f$' are not restricted to the physical quantities we recognize as time and frequency. Another 'domain pair' equivalent to time and frequency is space and wavenumber.

Basis functions can differ as long as they fulfil some orthogonality principle. In this view other transforms such as the Laplace and the wavelet transforms can be thought of as projections onto other 'basis coordinate systems'.

Transform techniques provide a very convenient way of describing general linear systems since operations in one domain are often simplified in the other domain. In the following, we give a list of the most important transform properties for the Fourier transform of which we will make use repeatedly ( $\Leftrightarrow$ indicates a transform pair, $x(t) \Leftrightarrow X(j\omega)$ ):

$$\bullet \textit{ Time shifting} — x(t-a) \Leftrightarrow X(j\omega)e^{-j\omega a} \text{ for } a > 0 \tag{2.11}$$

$$\bullet \textit{ Derivative} — \frac{d}{dt}x(t) \Leftrightarrow j\omega X(j\omega) \tag{2.12}$$

$$\bullet \textit{ Integration} — \int_{-\infty}^{t} x(\tau)d\tau \Leftrightarrow \frac{1}{j\omega}X(j\omega) \tag{2.13}$$

The properties of the Fourier transform with respect to the convolution of two signals are of special importance in the context of filtering. Convolution is the mathematical operation which allows us to calculate the output signal of a filter for a given input signal if we know a function which uniquely describes the filter's actions.

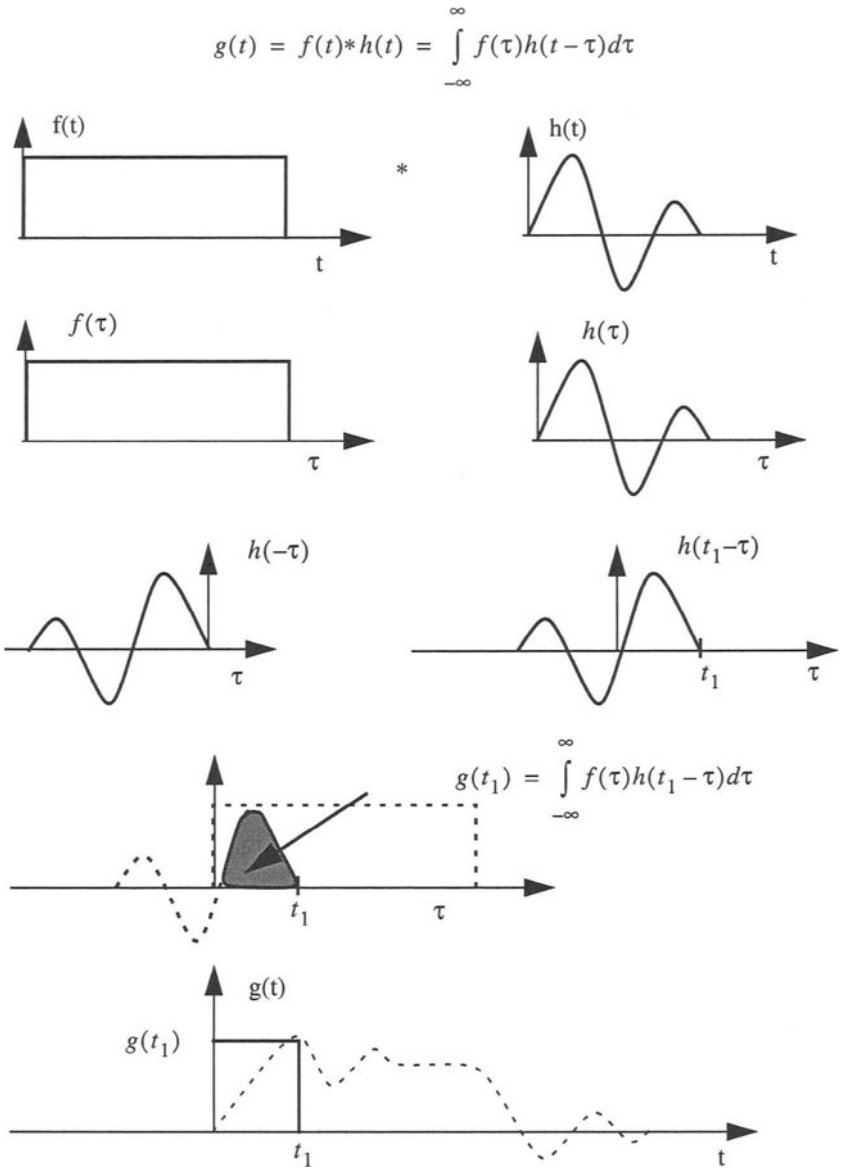



**Fig. 2.4** Graphical interpretation of the convolution operation.

The concept of characteristic filter functions - such as the impulse response function or the step response function - will be introduced in later chapters. Here, we simply give the definition of the convolution operation ( * ) on two functions f(t) and h(t):

$$g(t) = f(t)*h(t) = \int_{-\infty}^{\infty} f(\tau)h(t-\tau)d\tau \tag{2.14}$$

A graphical interpretation of the individual steps involved in this operation is displayed in Fig. 2.4. If the spectra corresponding to f(t) and h(t) are denoted $F(j\omega)$ and $H(j\omega)$, respectively, the transform pair for the convolution is given by the

$$\bullet\ \textit{Convolution Theorem} — \ f(t)*h(t) \Leftrightarrow F(j\omega)H(j\omega) \tag{2.15}$$

This theorem states that the calculation of the convolution integral in the time domain can be performed in the frequency domain by simply multiplying the corresponding spectra. It becomes especially important in the context of removing the effect of a filter (deconvolution) since it tells us that this can be done by simply dividing the spectrum of the signal by the spectrum describing the filter. In the time domain this operation has no equally simple equivalence. The proofs of these theorems are found for example in Kraniauskas (1992).

Returning to the problem of the RC filter, recall that we can obtain the output signal corresponding to an arbitrary input signal by considering the output signals for harmonic input signals $x(t) = A_i e^{j\omega t}$ and then superimposing the responses for the individual frequencies (making use of the linearity property equations (2.4) and (2.5) and the synthesis equation of the Fourier transform (2.10)). To solve equation (2.3) for a general harmonic signal, we propose the classical trial solution for the output signal:

$$y(t) = A_0 e^{j\omega t} \tag{2.16}$$

$$\dot{y}(t) = j\omega A_0 e^{j\omega t} \tag{2.17}$$

Inserting equations (2.16) and (2.17) into equation (2.3) we obtain:

$$A_0 e^{j\omega t}(RCj\omega + 1) = A_i e^{j\omega t} \tag{2.18}$$

and

$$\frac{A_0}{A_i} = \frac{1}{RCj\omega + 1} = T(j\omega) \tag{2.19}$$

$T(j\omega)$ is called the **frequency response function**. Strictly speaking, this is the so-called steady-state frequency response function for which we assume that the harmonic input signal has started long before the observation time. For a harmonic input signal which is 'turned on' shortly before the observation time, the output signal would consist of the superposition of the so-called transient response for which the amplitude decays with time and the steady-state response. In the present context, we will always assume that the transient response has already become insignificant and can be ignored.

For a harmonic input signal $x(t) = A_i e^{j\omega t}$ the corresponding output signal from (2.19) is $y(t) = A_o e^{j\omega t}$ with

$$A_0 = T(j\omega) \cdot A_i . \tag{2.20}$$

$T(j\omega)$ is a complex quantity. We can rewrite it in polar form by using the identities $1/(\alpha + j\beta) = \alpha/(\alpha^2 + \beta^2) - j\beta/(\alpha^2 + \beta^2)$ and $\Phi = \text{atan}(Im/Re)$. For $1/(\alpha + j\beta)$, $\Phi$ becomes $\text{atan}(-\beta/\alpha)$, thus

$$T(j\omega) = \frac{1}{\sqrt{1 + (RC\omega)^2}} e^{j\Phi(\omega)} \text{ with} \tag{2.21}$$

$$\Phi(\omega) = \text{atan}(-\omega RC) = -\text{atan}(\omega RC) . \tag{2.22}$$

For an harmonic input signal with frequency $\omega$, the phase shift due to the filter depends on frequency as well as on the product of capacitance and resistance.

*Problem 2.1* Calculate the output signal of the RC filter for a sinusoidal input signal $A_i \sin(\omega_0 t)$. Make use of Euler's formula ($\sin y = (e^{jy} - e^{-jy})/2j$) and equations (2.21) and (2.22).

The output of the filter for an harmonic input signal is again an harmonic signal with different amplitude and phase but with the same frequency. This property has its analogy in the eigenvector concept of linear algebra. Therefore, the values of the frequency response function are eigenvalues of the system with the harmonic functions being the corresponding eigenvectors (Fig. 2.5). In a linear transformation represented by a linear equation $\vec{x} \rightarrow A \cdot \vec{x}$, an eigenvector is transformed onto itself $\vec{x} \rightarrow \lambda \cdot \vec{x}$; hence its direction

does not change. If, by analogy, harmonic signals are taken to be 'eigenvectors' their frequency content does not change either.

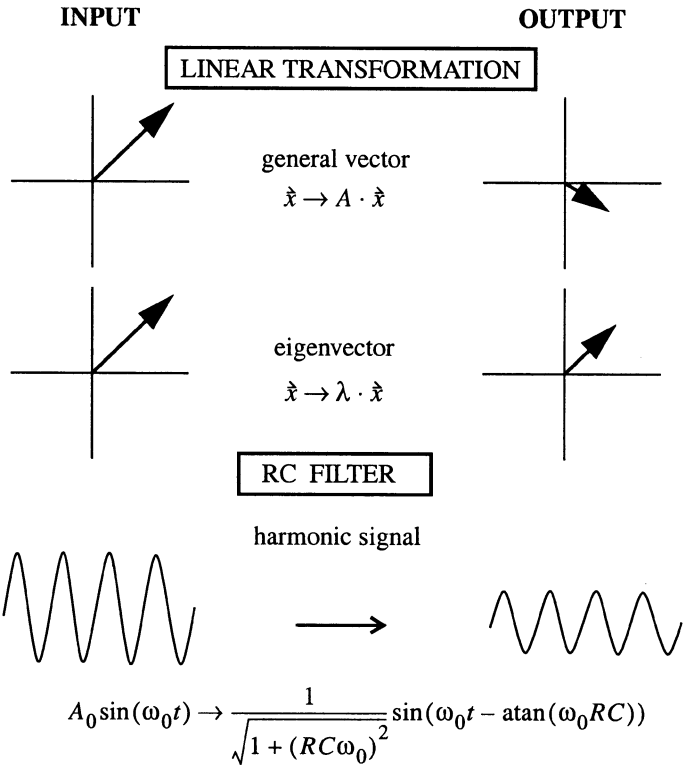



**Fig. 2.5** Frequency response function and the eigenvector/eigenvalue concept.

The frequency response of the system is valid for arbitrary input signals. In this case, we can visualize $A_i(j\omega)$ as an harmonic component of the input signal's Fourier spectrum $X(j\omega)$. $A_o(j\omega)$ is the corresponding harmonic component of the output signal's Fourier spectrum $Y(j\omega)$. Hence, the frequency response function relates the Fourier spectrum of the output signal $Y(j\omega)$ to the Fourier spectrum of the input signal $X(j\omega)$:

$$T(j\omega) = \frac{Y(j\omega)}{X(j\omega)} \tag{2.23}$$

In other words

> • *Definition* — The frequency response function $T(j\omega)$ is defined as the Fourier transform of the output signal divided by the Fourier transform of the input signal.

The properties of an arbitrary linear filter are uniquely described by its frequency response function. In general, the Fourier spectrum of a signal which has passed through a filter is obtained by multiplying (complex multiplication) the Fourier spectrum of the input signal by the frequency response function of the filter.

$$Y(j\omega) = T(j\omega) \cdot X(j\omega) . \tag{2.24}$$

The frequency response function is an extremely important tool in signal processing. It can be measured by comparing output and input signals to the system **without further knowledge of the physical processes going on inside the filter**. Once we know its frequency response function, we can predict a filter's output for any given input signal.

## 2.4 The transfer function and the Laplace transform

The frequency response function is closely related to the concept of the **transfer function**. To demonstrate, let us solve equation (2.3) again using the Laplace transformation. The bilateral Laplace transform of a function $f(t)$ is defined as (e.g. Papoulis, 1962):

$$\mathcal{L}[f(t)] = \int_{-\infty}^{\infty} f(t) e^{-st} dt \tag{2.25}$$

with the complex variable $s = \sigma + j\omega$. $\mathcal{L}[f(t)]$ is also a linear transform and will be written as $F(s)$. An important property of the Laplace transform in the context of solving differential equations is that the derivative in the time domain corresponds to a multiplication with $s$ in the Laplace domain which is also called the complex s-plane ( $\Leftrightarrow$ indicates a transform pair, $f(t) \Leftrightarrow F(s)$ ):

$$\bullet \ \frac{d}{dt} f(t) \Leftrightarrow sF(s) \text{ or equivalently } \mathcal{L}[\dot{f}(t)] = sF(s) \tag{2.26}$$

Transforming equation (2.3), we obtain:

$$RCsY(s) + Y(s) - X(s) = 0 \tag{2.27}$$

with $Y(s)$ and $X(s)$ being the Laplace transforms of $y(t)$ and $x(t)$, respectively.

> • *Definition* — The transfer function $T(s)$ is defined as the Laplace transform of the output signal divided by the Laplace transform of the input signal.

$$T(s) = \frac{Y(s)}{X(s)} = \frac{1}{1 + sRC} = \frac{1}{1 + s\tau} \tag{2.28}$$

If in equation (2.25), we set $s = j\omega$ and we follow through to equation (2.28) we obtain equation (2.19), the frequency response function. This is equivalent to the fact that the Fourier transform equals the Laplace transform evaluated along the imaginary axis of the s-plane.

If we look at equation (2.28), we can see that $T(s)$ grows to infinity for $s = -1/\tau$. It is said, that $T(s)$ has a **pole** at this location $s_p$. We will see in the following that the existence and the position of the pole at $s = s_p = -1/\tau$ are sufficient to describe most of the properties of the transfer function.

In signal processing literature, the term transfer function is sometimes used for the frequency response function as well. Although this is unfortunate, the meaning should become clear from the context. In the same way we used the Fourier transform to calculate the output spectrum using (2.24), we can calculate the Laplace transform of the output signal by multiplying the Laplace transform of the input signal with the transfer function:

$$Y(s) = T(s)X(s) \tag{2.29}$$

## 2.5 The impulse response function

In order to understand the concept of the impulse response function and its relationship to the frequency response function and to the transfer function, we introduce the so-called Dirac delta function $\delta(t)$. Mathematically, $\delta(t)$ is defined by the following equation (e.g. Papoulis, 1962)

$$\int_{-\infty}^{\infty} f(t)\delta(t-\tau)dt = f(\tau) \tag{2.30}$$

It defines the delta function by a convolution operation with a function $f(t)$. This convo-

lution yields the value of the function at that time ($\tau$) for which the argument of the delta function ($t-\tau$) vanishes. In other words the convolution of a function with the delta function yields the original function. This is called the sifting property of the delta function. Although $\delta(t)$ is not a function in the usual sense and does not exist other than in the context of a convolution integral, it is generally sufficient to think of it as a pulse of infinitesimal duration, infinite height and unit area. We can generate such an impulse by reducing the duration of a finite boxcar signal with unit area to zero duration as illustrated schematically in Fig. 2.6.

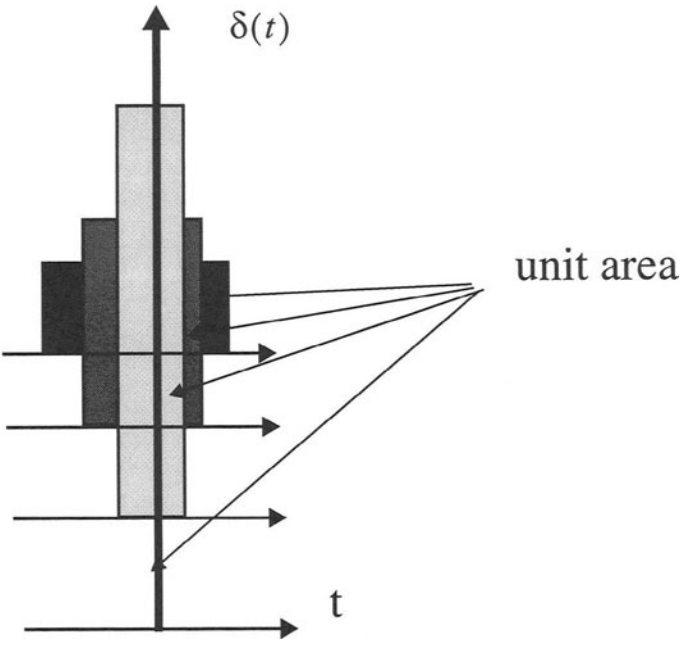



**Fig. 2.6** Generation of a delta function.

From the conditions of unit area and infinitesimal duration we obtain

$$\int_{-\infty}^{\infty} \delta(t)dt = 1 \tag{2.31}$$

and

$$\delta(t) = 0 \text{ for } t \neq 0. \tag{2.32}$$

Since $\int_{-\infty}^{t_1} \delta(t)dt = 0$ for $t_1 < 0$ and 1 for $t_1 \geq 0$, we can write

$$\int_{-\infty}^{t} \delta(\tau)d\tau = u(t) \tag{2.33}$$

with $u(t)$ being the unit step function ($u(t) = 0$ for $t < 0$ and 1 for $t \geq 0$). Hence, the Dirac delta function is the derivative of the unit step function

$$\delta(t) = \frac{du(t)}{dt} \tag{2.34}$$

The delta function has the important property that both its Fourier transform and Laplace transform are 1 as we can easily see from the definitions of the Fourier (equation (2.8)) and the Laplace transform (equation (2.25)), and the definition of the delta function (2.30) for $\tau = 0$:

$$\mathcal{F}\{\delta(t)\} = \int_{-\infty}^{\infty} \delta(t)e^{-j2\pi ft}\, dt = 1 \tag{2.35}$$

$$\mathcal{L}[\delta(t)] = \int_{-\infty}^{\infty} \delta(t)\, e^{-st}\, dt = 1 \tag{2.36}$$

Now we have the background necessary to deal quantitatively with the concept of the **impulse response function** $h(t)$ and its relationships to the frequency response function and the transfer function. The impulse response function is defined as the response of a filter to an impulsive (delta function) input signal.

Using (2.35) and (2.36) we can see from the definition of the frequency response function (2.23) and the definition of the transfer function (2.28) that the following relationships hold

• The frequency response function $T(j\omega)$ is the Fourier transform of the impulse response function.

and

• The transfer function $T(s)$ is the Laplace transform of the impulse response function.

To prove these statements we look at the transfer function $T(s)$ for $x(t) = \delta(t)$ and hence $X(s) = 1$. In this case, the output signal $y(t)$ becomes the impulse response function $h(t)$ with $H(s)$ being its Laplace transform.

$$T(s) = \frac{Y(s)}{X(s)} = \frac{Y(s)}{1} = H(s) \text{ for } x(t) = \delta(t). \tag{2.37}$$

The same argument can be made for the frequency response function

$$T(j\omega) = \frac{Y(j\omega)}{X(j\omega)} = \frac{Y(j\omega)}{1} = H(j\omega) \text{ for } x(t) = \delta(t). \tag{2.38}$$

From equation (2.24) we know that the Fourier spectrum of a filter output signal $Y(j\omega)$ is the product of the frequency response function $T(j\omega)$ with the Fourier spectrum of the input signal $X(j\omega)$. The convolution theorem (2.15) states that the multiplication of two spectra is equivalent to convolving the corresponding time functions. Equation (2.38) on the other hand tells us that the frequency response function and the impulse response function are a Fourier pair. Consequently, the filter process can mathematically be described not only by multiplying the Fourier spectrum of the input signal $X(j\omega)$ with the frequency response function $T(j\omega)$ to obtain the Fourier spectrum of a filter output signal $Y(j\omega)$ but also by convolving the input signal $x(t)$ with the impulse response function $h(t)$ to obtain the output signal $y(t)$.

An additional way to describe the filter process is to multiply the Laplace transform of the input signal $X(s)$ with the transfer function $T(s)$ to obtain the Laplace transform of the output signal $Y(s)$. Depending on the circumstances of each particular situation we can choose to apply a filter operation either in the 'time domain' or in one of the 'frequency domains'.

To determine the impulse response function for the RC filter, we rewrite its transfer function given in equation (2.28) as $F(s) = K/(s+a)$ with $K = 1/\tau$ and $a = 1/\tau$. With reference to one of the advanced texts on integral transforms (e.g. Kraniauskas, 1992), $F(s)$ is the Laplace transform of the signal $f(t) = K \cdot e^{-at}u(t)$. Here $u(t)$ is the unit step function defined by equation (2.34). Hence $f(t)$ is the impulse response function of the RC filter. This can be demonstrated as follows. We obtain the Laplace transform of $f(t)$ from (2.25) as

$$F(s) = K\int_{-\infty}^{\infty} e^{-at}e^{-st}u(t)dt = K\int_{0}^{\infty} e^{-(s+a)t}dt = -K\frac{e^{-(s+a)t}}{s+a}\Bigg|_{0}^{\infty} \qquad (2.39)$$

We can see right away that (2.39) exists only for Re{s+a} > 0 or Re{s} > Re{-a} where the result becomes $K/(s+a)$. Hence, replacing $K$ and $a$ by $1/\tau$ ($\tau = RC$) again, we see that the transfer function of the RC filter (2.28) is the Laplace transform of

$$y(t) = \frac{1}{\tau}e^{-\frac{1}{\tau}t} \qquad \text{for } t > 0 \qquad (2.40)$$

Therefore, equation (2.40) describes the impulse response of the RC filter.

The region where (2.39) exists is called *region of convergence* of the Laplace transform and in this case becomes the right half-plane with Re {s} > -1/$\tau$ as shown by the shaded region in Fig. 2.7.

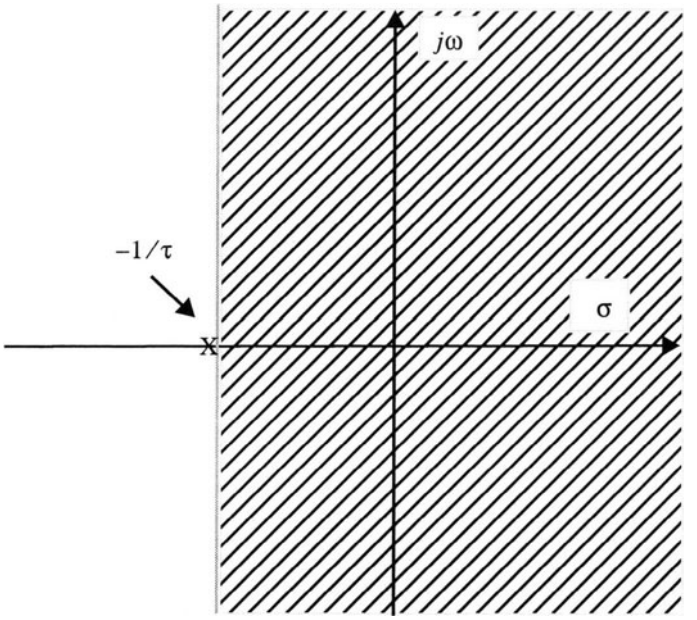



**Fig. 2.7** Region of convergence of (2.39). The pole location at $-1/\tau$ is marked by an X.

Since we are using the bilateral Laplace transform, it should be noted that the Laplace transform of $f(t) = -K \cdot e^{-at}u(-t)$, with $u(-t)$ being the time inverted unit step function:

$$F(s) = -K \int_{-\infty}^{\infty} e^{-at}e^{-st}u(-t)dt = -K \int_{-\infty}^{0} e^{-(s+a)t}dt = K\frac{e^{-(s+a)t}}{s+a}\Bigg|_{-\infty}^{0} \tag{2.41}$$

also comes out to be $K/(s+a)$. However, we see that it exists only for Re{s+a} < 0 or Re{s} < Re{-a} which corresponds to the shaded region in Fig. 2.8. Hence, **for the same transfer function we can obtain different impulse response functions!**

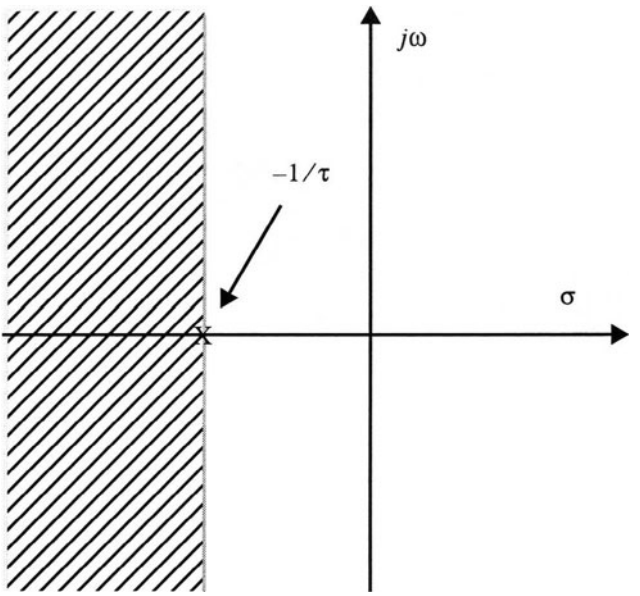



**Fig. 2.8** Region of convergence of (2.41). The pole location at $-1/\tau$ is marked by an X.

When we formally calculate the impulse response function from the transfer function, we technically calculate the inverse Laplace transform, defined as

$$\mathcal{L}^{-1}[F(s)] = f(t) = \frac{1}{2\pi j}\int_{\sigma - j\infty}^{\sigma + j\infty} F(s)e^{st}ds\,. \tag{2.42}$$

The path of integration must lie in the region of convergence. If we evaluate (2.42) on the imaginary axis, which means for $s = j\omega$, (2.42) becomes the equation of the inverse Fourier transform (equation (2.10)). Since for $s = j\omega$ the transfer function becomes the frequency response function, this means that the impulse response can either be calculated by inverse Fourier transform from the frequency response function or by inverse Laplace transform from the transfer function.

Depending on whether the region of convergence considered for the evaluation of (2.42) is a right half-plane or a left half-plane, we will get a right-sided or a left-sided impulse response function. For the example given, $F(s) = K/(s+a)$, the right-sided function $f(t) = K \cdot e^{-at}u(t)$ corresponds to a causal system ($f(t) = 0$ for $t < 0$) while the left-sided function $f(t) = -K \cdot e^{-at}u(-t)$ corresponds to an anticausal signal which vanishes for $t > 0$ and is physically not realizable.

### 2.5.1 THE CONDITION FOR STABILITY

The physically realizable impulse response of the RC filter $[h(t) = (1/\tau)e^{(-1/\tau)t}u(t)]$ has an exponential time dependence with the exponent $(-1/\tau)$ being exactly the location of the pole $(s_p)$ of the transfer function. As long as the pole is located in the left half of the complex s-plane, the causal impulse response $(|s_p|e^{s_p})$ will decay exponentially with time. However, if the pole is located within the right half-plane, the impulse response will become unstable (growing beyond bounds). This rule is also valid for more complicated transfer functions:

> For a causal system to be stable, all the poles of the transfer function must be located within the left half of the complex s-plane.

It should be noted however, that for anticausal signals the opposite is true. For a pole at $1/\tau$, the anticausal signal $y(-t) = (1/\tau)e^{(1/\tau)(-t)}u(-t)$ would be stable, although physically unrealizable.

## 2.6 The step response function

Although not as frequently used as the impulse response function, the response of a linear system to a unit step $u(t)$, the so-called **step response function**, is of considerable importance in seismological contexts. One reason for its importance is that experiments to calibrate seismic recording systems often use excitation signals which are step functions. It is easy to see that a step function can be generated simply by turning a voltage on or off. We will see how closely the step response function and the impulse response function are related and that they provide equivalent descriptions of a system. To show this, we will make use of some properties of the Laplace transform, which are reviewed below.

Since the Laplace transform can be seen as an extension of the Fourier transform, its transform properties can be intuitively understood from the transform properties of the Fourier transform which we have discussed on page 16. Again $\Leftrightarrow$ indicates a transform pair, $x(t) \Leftrightarrow X(s)$.

In addition to the time derivative property (2.12)

$$\bullet \textit{ Derivative} — \frac{d}{dt}x(t) \Leftrightarrow sX(s)$$

which was already introduced above we will use of the following transform pairs:

$$\bullet \textit{ Unit step} — u(t) \Leftrightarrow \frac{1}{s} \text{ for } Re\{s\} > 0 \tag{2.43}$$

$$\bullet \textit{ Integration} — \int_{-\infty}^{t} x(\tau)d\tau \Leftrightarrow \frac{1}{s}X(s) \text{ for } max(\alpha, 0) < Re\{s\} < \beta \tag{2.44}$$

Since $1/s$ has a pole at the origin, the lower limit of the region of convergence is restricted to the right half s-plane even if the region of convergence for $X(s)$ defined by $\alpha < Re\{s\} < \beta$ is larger.

$$\bullet \textit{ Convolution}— x(t) * h(t) \Leftrightarrow X(s) \cdot H(s)$$
$$\text{for } max(\alpha_1, \alpha_2) < Re\{s\} < min(\beta_1, \beta_2) \tag{2.45}$$

Here $(\alpha_1, \beta_1)$ and $(\alpha_2, \beta_2)$ define the regions of convergence for $X(s)$ and $H(s)$, respectively. Proofs of these properties can be found for example in Kraniauskas (1992). From (2.43) - (2.45) we can see that convolving a function with the unit step function, hence multiplying its Laplace transform with $1/s$, is equivalent to integration.

The system response to a unit step is obtained easily from the definition of the transfer function $T(s) = \dfrac{Y(s)}{X(s)}$ (equation (2.28)) if we let $x(t) = u(t)$.

From (2.43) we obtain for $X(s) = U(s) = \dfrac{1}{s}$ and

$$T(s) = \frac{Y(s)}{1/s} \text{ or} \tag{2.46}$$

$$Y(s) = \frac{T(s)}{s} \tag{2.47}$$

From (2.44) we see that this corresponds to integrating the impulse response function. Hence the step response function $a(t)$ is obtained as

$$a(t) = \int_{-\infty}^{t} h(\tau)d\tau \tag{2.48}$$

and

$$h(t) = \frac{d}{dt}a(t) \tag{2.49}$$

We obtain the following rule

> • The step response function $a(t)$ and the impulse response function $h(t)$ are equivalent descriptions of a system and can be obtained from each other via integration or differentiation, respectively.

Although we can calculate the impulse response function of a system once the step response function is obtained, it is also possible to directly express the response of a filter $g(t)$ to an arbitrary input $f(t)$ in terms of the step response function $a(t)$ using

$$g(t) = A(0)f(-\infty) + \int_{-\infty}^{t} f'(\tau)a(t-\tau)d\tau \tag{2.50}$$

Here $A(0)$ is the value of the modulus of the frequency response function at zero frequency and $f(-\infty)$ is the value of the input function at $t = -\infty$. The proof of (2.50) can be found in Papoulis (1962).

## 2.7 The frequency response function and the pole position

Given the pole location of the transfer function on the complex s-plane we can determine the frequency response function of the system using a simple graphical method (Fig. 2.9).

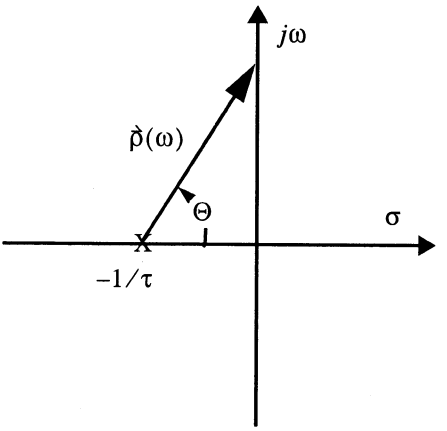



**Fig. 2.9** Representation of the RC filter in the s-plane. The pole location at $-1/\tau$ is marked by an X.

For $s = j\omega$, $s$ moves along the imaginary axis. Rewriting equation (2.28) we get:

$$T(s) = \frac{1}{1+s\tau} = \frac{1}{\tau}\left[\frac{1}{\frac{1}{\tau}+s}\right] \tag{2.51}$$

and with $s = j\omega$

$$T(j\omega) = \frac{1}{\tau}\left[\frac{1}{\frac{1}{\tau}+j\omega}\right] \tag{2.52}$$

Written as a complex number, $1/\tau + j\omega$ represents the pole vector $\hat{p}(\omega)$ which points from the pole position towards the actual frequency on the imaginary axis. Using $\hat{p}(\omega)$ the length of the vector and polar coordinates we obtain for the frequency response function:

$$T(j\omega) = \frac{1}{\tau}\left[\frac{1}{|\hat{p}(\omega)|e^{j\theta(\omega)}}\right] = \frac{1}{\tau}\left[\frac{1}{|\hat{p}(\omega)|}e^{-j\theta(\omega)}\right] = |T(j\omega)|e^{j\Phi(\omega)} \tag{2.53}$$

For the given example, the amplitude value of the frequency response function for frequency $\omega$ is inversely proportional to the length of the pole vector $\vec{p}(\omega)$ from the pole location to the point $j\omega$ on the imaginary axis. The phase angle equals the negative angle between $\vec{p}(\omega)$ and the real axis.

*Problem 2.2* Determine graphically the modulus of the frequency response function for a RC filter with $R = 4.0\Omega$ and $C = 1.25F/2\pi = 0.1989495F$ $(1\Omega = 1(V/A),$ $1F = 1A\text{sec}/V)$. Where is the pole position in the s-plane? For the plot use frequencies between 0 and 5 Hz.

The graphical method of determining the frequency response function is very instructive, since it provides a quick look at the system's properties in the frequency domain. However, for a quantitative analysis we should determine $T(j\omega)$ by directly evaluating equation (2.52) for different frequency values. A numerical solution to Problem 2.2 is shown in Fig. 2.10.

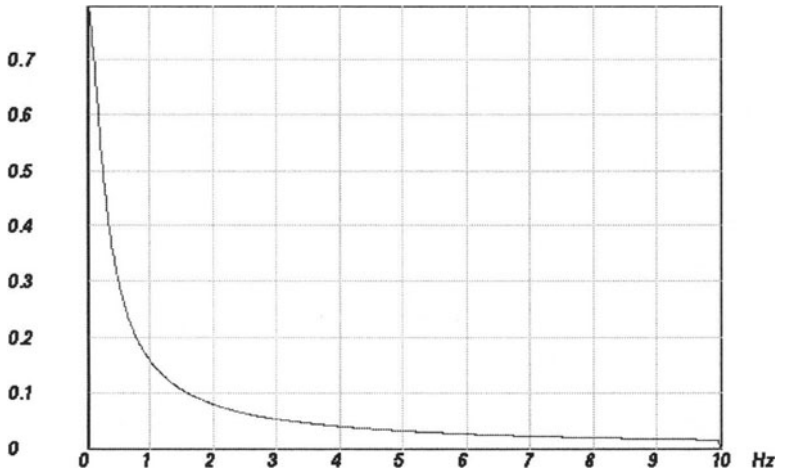



**Fig. 2.10** Frequency response function (amplitude only) of the RC filter of Problem 2.2.

We see that for high frequencies, the amplitude values of the frequency response function decrease continuously. We could have guessed this already (or known from physics class), since for low frequencies the capacitor will act like an resistor of infinite resistance, while for high frequencies it will act like a short circuit. Therefore, the high frequency components of the input signal will not make it to the output signal. In other words, the circuit acts as a lowpass filter.

*Problem 2.3* Calculate the frequency response of the RC filter from Problem 2.2 using the Digital Seismology Tutor (DST).

We can learn some more about the effect of the pole on the frequency response function by displaying the same amplitude spectrum again on a log-log plot (Fig. 2.11).

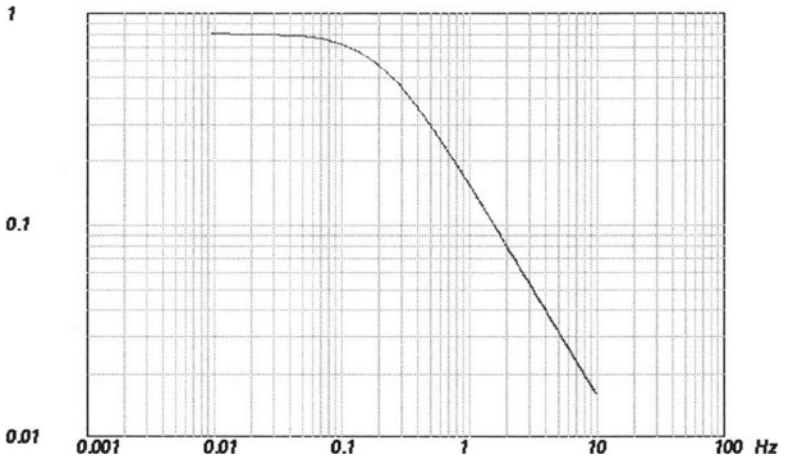



**Fig. 2.11** Same plot as Fig. 2.10 only on a log-log scale.

Here, we can see that the amplitude portion of the frequency response function can be approximated in the high- and low frequency limit roughly by two straight lines. These lines intersect at a frequency of 0.2 Hz which is called the corner frequency of the filter. For the RC filter, it is equal to $1/(2\pi \cdot RC) = 1/(5\text{sec})$ which is also the exponential term in the impulse response.

We can understand this more quantitatively by going back to equation (2.52) and concentrating on the amplitude portion of the frequency response function.

$$|T(j\omega)| = \frac{1}{\tau}\left[\frac{1}{\left|\frac{1}{\tau} + j\omega\right|}\right] = \frac{1}{\sqrt{1+\omega^2\tau^2}} \tag{2.54}$$

$|-1/\tau|$ is the distance of the pole position from the origin of the s-plane. If we define this as $\omega_c$, then we can write

$$|T(j\omega)| = \frac{1}{\sqrt{1+\frac{\omega^2}{\omega_c^2}}} \tag{2.55}$$

Let us consider the behaviour of this function as $\omega \to 0$ and for $\omega \gg \omega_c$. For $\omega \to 0$, $|T(j\omega)| \to 1 = \text{const}$, while for $\omega \gg \omega_c$ we can easily see that $|T(j\omega)| \approx \omega^{-1}$. Thus, the slope of the frequency response function (amplitude portion) in a log-log plot changes from being 0 for frequencies much smaller than $\omega_c$ to -1 for frequencies much larger than $\omega_c$. At the corner frequency $\omega_c$ the amplitude of the frequency response function $|T(j\omega)|$ is $1/\sqrt{2}$. If we measure the slope from the amplitude ratio $A(\omega_1)$ and $A(\omega_2)$ at frequencies $\omega_1$ and $\omega_2$ we obtain $Slope_{\log-\log} = (\log_{10}A(\omega_2)-\log_{10}A(\omega_1))/(\log_{10}(\omega_2)-\log_{10}(\omega_1))$ which equals $\log_{10}(A(\omega_2)/A(\omega_1))/\log_{10}(\omega_2/\omega_1)$.

If the amplitude ratio is expressed in *dB* $(20 \cdot \log_{10}(\text{amplitude ratio}))$, the slope becomes $Slope_{dB/\Delta\omega} = 20 \cdot \log_{10}(A(\omega_2)/A(\omega_1))/\log_{10}(\omega_2/\omega_1)$ in $[dB/\Delta\omega]$. For a frequency response function $|T(j\omega)| \approx \omega^{-1}$ the amplitude decreases by a factor of 10 over a full decade $(\omega_2 = 10 \cdot \omega_1)$. Therefore the slope in this case becomes $Slope_{dB/dec} = 20 \cdot \log_{10}(0.1)/\log_{10}(10) = -20$ $[dB/decade]$ or following the same argument -6 *dB/octave* $(\omega_2 = 2 \cdot \omega_1)$. We can state the general rule:

> A single pole in the transfer function causes the slope of the amplitude portion of the frequency response function in a log-log plot to decrease by 20 *dB/decade* or 6 *dB/octave*, for frequencies larger than the corner frequency $\omega_c$.

We can also apply this rule the other way round. From the decrease of the slope of the amplitude portion of a given frequency response function at a corner frequency $\omega_c$, we can expect a pole to exist which must be located at a distance $|\omega_c|$ away from the origin of the s-plane. We will see later how this rule can be extended to zeros of the transfer function and to more complicated transfer functions as well.

## 2.8 The difference equation

Yet another method to represent systems, extensively used in processing discrete data, is by difference equations. For the RC circuit, we arrive at such a representation by approximating the derivative at time $t$ in equation (2.3) by a finite difference at time $nT$, with the discretization interval being $T$

$$\dot{y}(t) = \frac{dy(t)}{dt} = \frac{dy(nT)}{dt} \overset{\text{for small } \Delta t}{\approx} \frac{\Delta y(nT)}{\Delta t} = \frac{y(nT)-y((n-1)T)}{T} \tag{2.56}$$

Equation (2.3) becomes

$$RC\frac{y(nT)-y((n-1)T)}{T} + y(nT) = x(nT)\,. \tag{2.57}$$

Writing $y(nT)$ as $y(n)$ this leads to

$$y(n) = b_0 x(n) - a_1 y(n-1) \tag{2.58}$$

with

$$b_0 = \frac{\frac{T}{RC}}{1+\frac{T}{RC}}$$
$$a_1 = -\frac{1}{1+\frac{T}{RC}} \tag{2.59}$$

The output signal $y(n)$ at time $nT$ depends on the value of the input signal $x(n)$ at time $nT$ as well as on the value of the output signal at time $(n-1)T$. Equations (2.58) and (2.59) can directly be used for the numerical implementation of the RC filter as a recursive filter. It should be emphasized that the 'discretization' of equation (2.3) is only valid for $T \to 0$. Furthermore, it introduces the effect of aliasing which is discussed in chapter 5.

*Problem 2.4* Let us end this chapter by considering an example directly related to our daily life. Consider a savings account with a monthly interest rate of $\alpha$ percent. The money which is deposited at time $t = nT$ is supposed to be $x(nT)$. $y(nT)$ represents

the money in the account at time $nT$ (before the deposit of $x(nT)$ is made), and $y(nT+T)$ is the money one sample (1 month) later. Determine the difference- and differential equations of the system using the forward difference ($\dot{y}(t) \approx \frac{y(nT+T)-y(nT)}{T}$). Start out with the balance at time $t = nT+T$ which can be written as $y(nT+T) = y(nT) + \alpha y(nT) + \alpha x(nT) + x(nT)$. Calculate the transfer function using Laplace transform (use equation (2.26)). Is the system stable? Could we use an RC filter to simulate the savings account?

## 2.9 Review RC filter

The central theme of this chapter was to study the behaviour of a simple electric RC circuit. We introduced the term filter or system as a device or algorithm which changes some input signal into an output signal. We saw that the RC filter is an example for a **linear, time invariant (LTI) system**, which could be described by a linear differential equation. From the solution of the **differential equation** for harmonic input we obtained the result that the output is again a harmonic signal. We introduced the concept of the **frequency response function** as the Fourier transform of the output signal divided by the Fourier transform of the input signal. The frequency response function was seen to have important properties:

- The values of the frequency response function are the eigenvalues of the system.

- Knowing the frequency response function, we can calculate the output of the filter to arbitrary input signals by multiplying the Fourier transform of the input signal with the frequency response function.

- The frequency response function is the Fourier transform of the impulse response. Knowing the impulse response function, we can calculate the output of the filter to arbitrary input signals by convolving the input signal with the impulse response function.

We then introduced the concept of the **transfer function** as an even more general concept to describe a system as the Laplace transform of the output signal divided by the Laplace transform of the input signal. The transfer function can also be seen as the Laplace transform of the impulse response function.The frequency response function can be derived from the transfer function by letting $s = j\omega$. We found that the transfer function of the RC circuit has a pole at the location $-1/RC$ (on the negative real axis of the s-plane). We also found that the (causal) impulse response of a system with a single pole is proportional to an exponential function $e^{s_p t}$ with $s_p$ being the location of the pole. Therefore the causal system can only be **stable** if the pole is located within the left half-plane of the s-plane. Next, we introduced the step response function as yet another way

to express the action of a filter. It was shown that the step response function and the impulse response function are closely related and can be obtained from each other by integration and differentiation, respectively.

We found a way to graphically determine the frequency response function given the pole position in the s-plane. By analysing the frequency response function in a log-log plot, we derived the rule that a pole in the transfer function causes a change of the slope of the frequency response function at a frequency $\omega_c$ by 20 dB/decade, while $\omega_c$ is the distance of the pole from the origin of the s-plane. We finally approximated the **differential equation** of the RC circuit by its difference equation which can be solved iteratively.

CHAPTER 3

# General linear time invariant systems

## 3.1 Generalization of concepts

We studied the simple RC circuit in great detail because the concepts we used for its analysis are also valid for more complicated systems. In the following we will make the transition to general systems with only the restrictions of linearity and time invariance (LTI system). By the end of this chapter we will have progressed to frequency response functions relevant to seismological systems.

We begin generalizing our tools to accommodate more complex systems. Table 3.1 shows how the concepts we have been using to describe the RC-filter can be formally extended to describe general LTI systems. For example, if we rewrite the differential equation for the RC filter (2.3):

$$RC\dot{y}(t) + y(t) - x(t) = \alpha_1 \frac{d}{dt} y(t) + \alpha_0 y(t) + \beta_0 x(t) = 0 \tag{3.1}$$

we see it is a special case of an Nth order LTI system

$$\sum_{k=0}^{N} \alpha_k \frac{d^k}{dt} y(t) + \sum_{k=0}^{L} \beta_k \frac{d^k}{dt} x(t) = 0 \tag{3.2}$$

In chapter 2.5 we have seen that for a system with a single pole, the impulse response could be left-sided or right-sided depending on which region of convergence we considered. For a general LTI system, the regions of convergence consist of bands parallel to the imaginary ( $j\omega$ ) axis which do not contain any poles of the transfer function.

**Table 3.1** Correspondences between the RC filter (1st order system) and a general Nth order LTI system. For the RC filter $\alpha_1 = RC$, $\alpha_0 = 1$, $\beta_0 = -1$.

| Concept | RC filter | General system |
|---|---|---|
| Differential equation | $\alpha_1 \frac{d}{dt} y(t) + \alpha_0 y(t) + \beta_0 x(t) = 0$ | $\sum_{k=0}^{N} \alpha_k \frac{d^k}{dt} y(t) + \sum_{k=0}^{L} \beta_k \frac{d^k}{dt} x(t) = 0$ |
| Transfer function | $T(s) = \frac{-\beta_0}{\alpha_0 + \alpha_1 s}$ | $T(s) = \frac{\beta_0 + \beta_1 s + \beta_2 s^2 + \ldots + \beta_L s^L}{\alpha_0 + \alpha_1 s + \alpha_2 s^2 + \ldots + \alpha_N s^N}$ |
| Frequency response function | $T(j\omega) = \frac{-\beta_0}{\alpha_0 + \alpha_1 j\omega}$ | $T(j\omega) = \frac{\beta_0 + \beta_1 (j\omega) + \beta_2 (j\omega)^2 + \ldots + \beta_L (j\omega)^L}{\alpha_0 + \alpha_1 (j\omega) + \alpha_2 (j\omega)^2 + \ldots + \alpha_N (j\omega)^N}$ |
| Poles and zeros | A single pole at $-1/\alpha_1$, the root of the denominator polynomial $T(s) = \frac{-\beta_0}{\alpha_1 \cdot (s - s_{p1})}$ or $T(j\omega) = \frac{-\beta_0}{\alpha_1 (j\omega - s_{p1})}$ | N poles at the roots of the denominator polynomial, L zeros at the roots of the numerator polynomial $T(s) = \frac{-\beta_L \cdot \prod_{k=1}^{L} (s - s_{0k})}{\alpha_N \cdot \prod_{k=1}^{N} (s - s_{pk})}$ or $T(j\omega) = \frac{-\beta_L \cdot \prod_{k=1}^{L} (j\omega - s_{0k})}{\alpha_N \cdot \prod_{k=1}^{N} (j\omega - s_{pk})}$ |
| Difference equation | $y(n) = -a_1 y(n-1) + b_0 x(n)$ | $y(n) = -\sum_{k=1}^{N} a_k y(n-k) + \sum_{k=0}^{L} b_k x(n-k)$ |

For a right-sided signal, the region of convergence is always a right half-plane, while for a left-sided signal it is a left half-plane. For a two-sided signal it consists of a finite band. For the general LTI systems which will be considered here, the transfer functions are always rational functions. In this case the regions of convergence are always bounded by poles. Finally, for a signal of finite duration the region of convergence is always the complete s-plane. This will be an issue of importance later. Fig. 3.1 shows the different types of convergence regions for a general LTI system for which a Laplace transform exists and the corresponding types of stable impulse response functions (IR). If a system does not belong to any of these four classes, it does not have a Laplace transform.

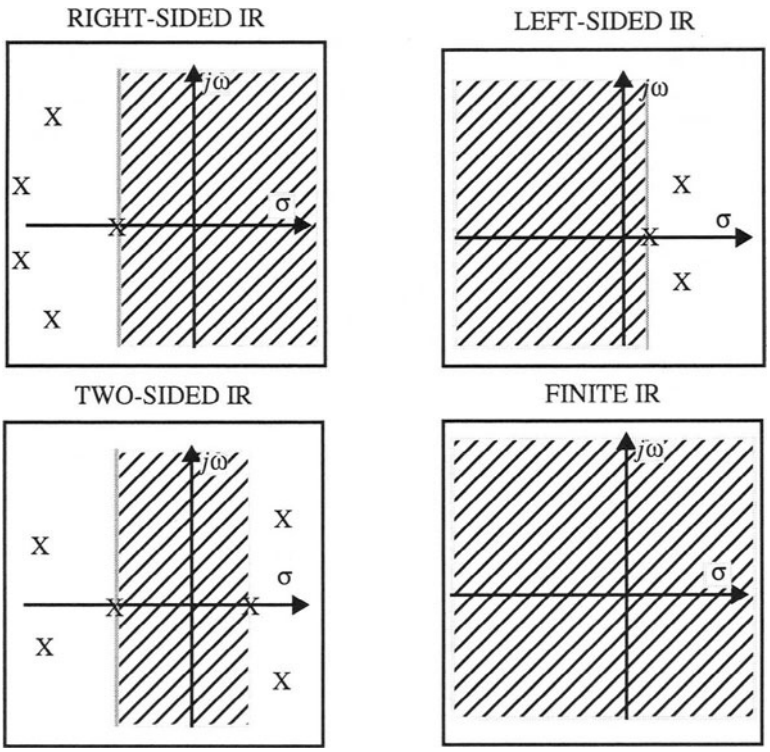



**Fig. 3.1** Types of convergence regions for general LTI systems with existing Laplace transform and the corresponding types of stable (infinite duration) impulse response (IR) functions.

When we calculate an impulse response function for a given pole-zero distribution numerically, we commonly use the inverse Fourier transform (actually the discrete Fourier transform, but we can ignore the differences for the following argument). Thus the path of integration is fixed to the imaginary axis. The impulse response will therefore be one for which the region of convergence contains the imaginary axis.

*Problem 3.1* System with two poles. Consider three different cases. a) Put both poles at (-1.2566, 0). b) Put one pole at location (-1.2566, 0) and the other one at (1.2566, 0). c) Put both poles at (1.2566, 0). For the input signal, use a spike at the center position of the window (for DST an internal sampling frequency of 100 Hz and a window length of 2048 points works well). What types of impulse response functions do you expect in each case? Will the frequency response functions be different? What changes do you expect for the frequency response functions with respect to the system in Problem 2.3?

### 3.2 Graphical estimation of the frequency response function

We see that the transition from the RC filter to the Nth order LTI system does not require major changes in the concepts of system analysis, only several extensions which are straightforward. The only new aspect is the occurrence of zeros in addition to poles. As it turns out, they can be treated in a very similar way to poles. Let us assume a system with a pole and a zero on the real axis of the s-plane (Fig. 3.2)

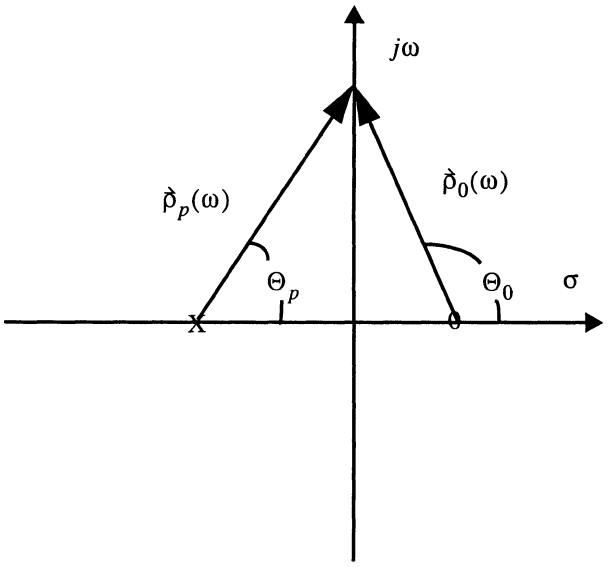



**Fig. 3.2** Complex s-plane representation of a system with a single pole and zero. The pole and zero locations are marked by a X and a 0, respectively.

The transfer function for this case becomes:

$$T(s) = \frac{s - s_0}{s - s_p} \tag{3.3}$$

with $s_0$ and $s_p$ being the position of the zero and the pole, respectively. For the frequency response function ($s = j\omega$) we get

$$T(j\omega) = \frac{j\omega - s_0}{j\omega - s_p} \tag{3.4}$$

Written as complex numbers, $j\omega - s_p$ and $j\omega - s_0$ represent the vectors $\vec{\rho}_p(\omega)$ and $\vec{\rho}_0(\omega)$ which point from the pole position and the zero position, respectively, towards the actual frequency on the imaginary axis. The frequency response function is then:

$$T(j\omega) = \left|\vec{\rho}_o(\omega)\right| e^{j\theta_0} \cdot \frac{1}{\left|\vec{\rho}_p(\omega)\right|} e^{-j\theta_p} \tag{3.5}$$

The amplitude value of the frequency response function for frequency $\omega$ is equal to the length of the vector $\vec{\rho}_0(\omega)$ from the position of the zero to the point $j\omega$ on the imaginary axis divided by the length of the vector $\vec{\rho}_p(\omega)$ from the position of the pole to the point $j\omega$ on the imaginary axis. The phase angle equals the angle between $\vec{\rho}_o(\omega)$ and the real axis minus the angle between $\vec{\rho}_p(\omega)$ and the real axis.

We can extrapolate this example to obtain a graphical method for determining the frequency response function of an arbitrary LTI system:

> The amplitude part of the frequency response function of an arbitrary LTI system can be determined graphically by multiplying the lengths of the vectors from the zero locations in the s-plane to the point $j\omega$ on the imaginary axis and then dividing by the product of the lengths of all vectors from pole locations to the point $j\omega$ on the imaginary axis. Likewise, to determine the phase part, the phase angles for the vectors from the zero locations in the s-plane to the point $j\omega$ on the imaginary axis are added together. Then, the phase angles of all the vectors from pole locations to the point $j\omega$ on the imaginary axis are subtracted from this sum.

*Problem 3.2* Use the argument given above to determine the frequency response for Problem 2.3 if you add a zero at position (1.2566, 0)?

### 3.3 The phase properties of general LTI system

In the preceding discussion it has been shown that for the amplitude part of the frequency response we must consider only the lengths of the pole- and zero-vectors, $\vec{\rho}_p(\omega)$ and $\vec{\rho}_o(\omega)$, respectively. For the example shown in Fig. 3.3 this means that because the zeros in a) and b) appear as mirror images with respect to the imaginary axis, both systems will have the same amplitude response. For the phase response we can easily see that the angle between $\vec{\rho}_o(\omega)$ and the real axis will always be greater for a zero in the right half-plane (Fig. 3.3b) than for a zero in the left half-plane (Fig. 3.3a).

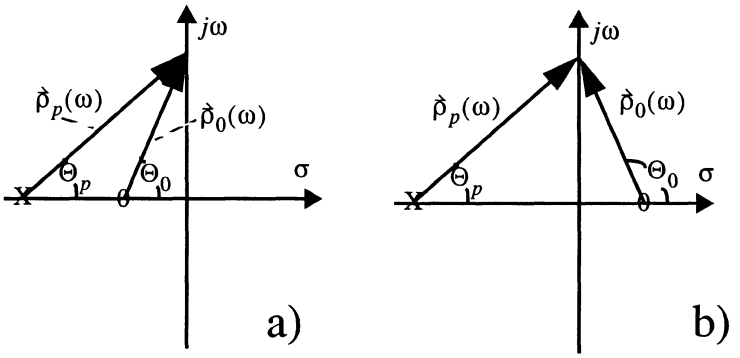



**Fig. 3.3** Complex s-plane representation of two systems with a single pole and zero having the same amplitude response. The pole and zero locations are marked by a X and a 0, respectively.

For general LTI systems, zeros in the right half-plane will always have larger phase response contributions than zeros in the left half-plane. For a given amplitude response, the phase response will have the smallest possible values if all zeros are located in the left half-plane. This leads us quite naturally to the concept of minimum and maximum phase:

> • *Definition* — A causal, stable system (no poles in the right half-plane) is *minimum phase* provided it has no zeros in the right half-plane. It is *maximum phase* if all its zeros are in the right half-plane.

It should be emphasized that the causality and stability conditions are prerequisites for the minimum phase property. Hence the transfer function of a minimum phase system has no singularities in the right half-plane. Minimum phase systems have a number of desirable properties which are especially important in the context of digital filters. Systems which are neither minimum phase nor maximum phase are called *mixed phase*. If a filter performs no phase distortion but causes a constant time shift for all frequencies, its phase response must be directly proportional to frequency. This can be understood from the shifting property of the Fourier transform (cf. equation (2.11)). These types of filters are called l*inear phase*. Filters for which the phase response is zero for all frequencies are called *zero phase* filters. They can be created implicitly by filtering the same signal twice in opposite directions using the same filter, thus cancelling the phase responses.

*Problem 3.3* How can the following two statements be proven for a general LTI system? a) If a system is minimum phase it will always have a stable and causal inverse filter. b) Any mixed phase system can be seen as a convolution of a minimum phase system and a filter which only changes the phase response but leaves the amplitude response as is (allpass filter).

*Problem 3.4* How can we change the two-sided impulse response from Problem 3.1b into a right-sided one without changing the amplitude response? Keyword: allpass filter.

### 3.4 The interpretation of the frequency response function

From the interpretation of the RC circuit we concluded that a single pole in the transfer function causes the slope of the amplitude portion of the frequency response function in a log-log plot to decrease by 20 dB/decade (6 dB/octave). The transition occurs at the frequency $\omega_c$ which is equal to the modulus of the pole position. If we take the inverse of the transfer function, a single pole will become a single zero and we can conclude likewise that a single zero in the transfer function causes the slope of the amplitude portion of the frequency response function in a log-log plot to **increase** by 20 dB/decade or 6 dB/octave. The transition takes place at a frequency $\omega_c$ which is equal to the modulus of the zero position.

*Problem 3.5* Consider a system with a pole and a zero on the real axis of the s-plane. Let the pole position be (-6.28318, 0), and the zero position (.628318, 0). What is the contribution of the zero to the frequency response function?

Using the general rule above, we can interpret the shape of the amplitude part of the frequency response of a general LTI system directly in terms of the locations of poles and zeros. A multiple pole or zero will contribute a slope change by a multiple of 20 dB/decade.

Before we end the discussion of the effects of poles and zeros on the transfer function, there is one more point to discuss. So far, we have only considered poles and zeros on the real axis of the s-plane. On the other hand, from the definition of the corner frequency $\omega_c$, we have seen that it depends only on the distance of the singularity from the origin of the s-plane. In other words, all poles located on a circle around the origin of the s-plane will produce the same corner frequency $\omega_c$ in the frequency response function. So what changes if we move the position of our single pole to another position along a circle?

But are we allowed to just shift a single pole along a circle? We have assumed so far that all our systems are linear time invariant systems. This means the systems could be described by linear differential equations. As long as the coefficients of the differential equations are real quantities, the coefficients of the polynomials making up the transfer function are real as well. The roots of a polynomial with real coefficients, however, can only be real or appear in complex conjugate pairs. So to answer our question: No, we cannot just shift a pole from the real axis. Poles and zeros appearing off the real axis of the s-plane must appear in conjugate complex pairs.

*Problem 3.6* Move the pole position of a double pole at (-1.2566, 0) in steps of $15^o$ (up to $75^o$, and $85^o$) along a circle centered at the origin and passing through the original double pole. Calculate the impulse response functions and the amplitude portions of the frequency response functions.

We can conclude that moving the poles along a circle around the origin of the s-plane does not change the spectral roll-off away from the corner frequency. However, it strongly changes the behaviour of the filter at the corner frequency. The closer the poles move towards the imaginary axis, the more does the impulse response show 'resonance' effects.

*Problem 3.7* Use the pole-zero approach to design a notch filter suppressing unwanted frequencies at 6.25 Hz. What kind of singularities do we need? How can we make use of the result of Problem 3.6?

*Problem 3.8* From the shape of the frequency response function in Fig. 3.4, determine the poles and zeros of the corresponding transfer function.

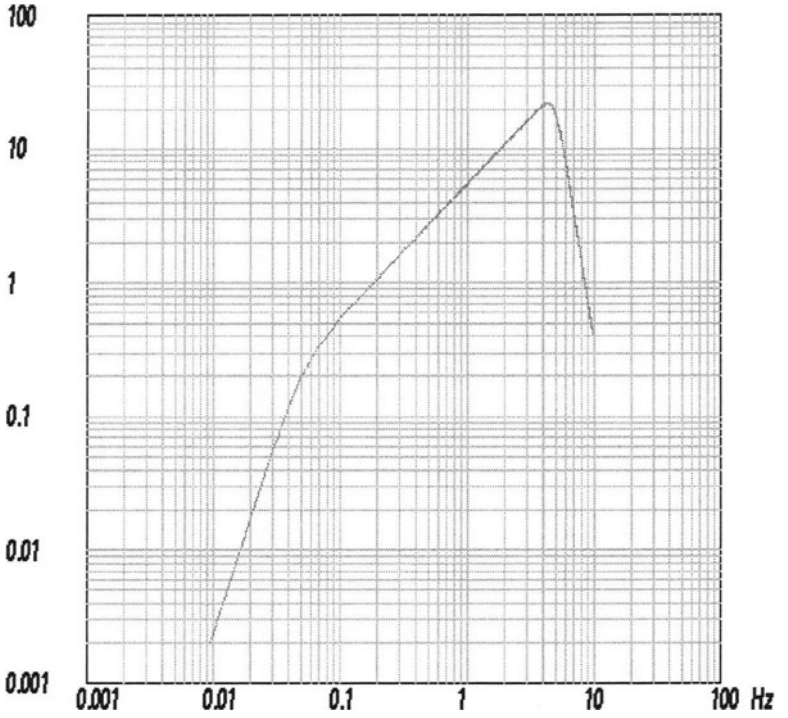



**Fig. 3.4** Frequency response function (amplitude) with an 'unknown' pole - zero distribution.

CHAPTER 4

# The seismometer

So far we have encountered several methods for describing linear time invariant systems. We have learned to design and analyse simple filters in terms of the poles and zeros of the transfer function. We have also seen some relationships between the different approaches.

Now we will apply these concepts to the classical pendulum seismometer. By analysing the frequency response and transfer functions of the seismometer we will learn how and why seismic signals recorded by a seismometer differ from the true ground motion. We will then be a step closer towards our final goal: removing the effects of the recording process from recorded seismograms.

To understand how seismic signals are altered by a seismometer and how the output signal relates to the true ground motion, we will use concepts from the last chapter to describe a seismometer as a linear time invariant system. We will restrict ourselves to a simple vertical pendulum seismometer such as is sketched in Fig. 4.1. It consists of three different elements: A mass, a spring and a damping mechanism usually taken to be a dashpot. The mass is connected to a frame by a spring and a dashpot and the frame is fixed relative to the ground. We can describe the motion of the mass relative to this fixed frame by the mass position $x_r(t)$. $\dot{x}_r(t)$ and $\ddot{x}_r(t)$ are its velocity and acceleration with respect to the frame. The position of the mass relative to the inertial system will be denoted $u_m(t)$. The movements of the mass are, of course, controlled by the sum of all forces acting on it:

• *The inertia of the mass* — The strength of the inertial force is proportional to the acceleration of the mass $m$. It is directed in the opposite direction of the motion (for example downwards if the ground moves upwards). With respect to the inertial reference frame denoted as $u$ it is given as

$$f_i = -m\ddot{u}_m(t) \tag{4.1}$$

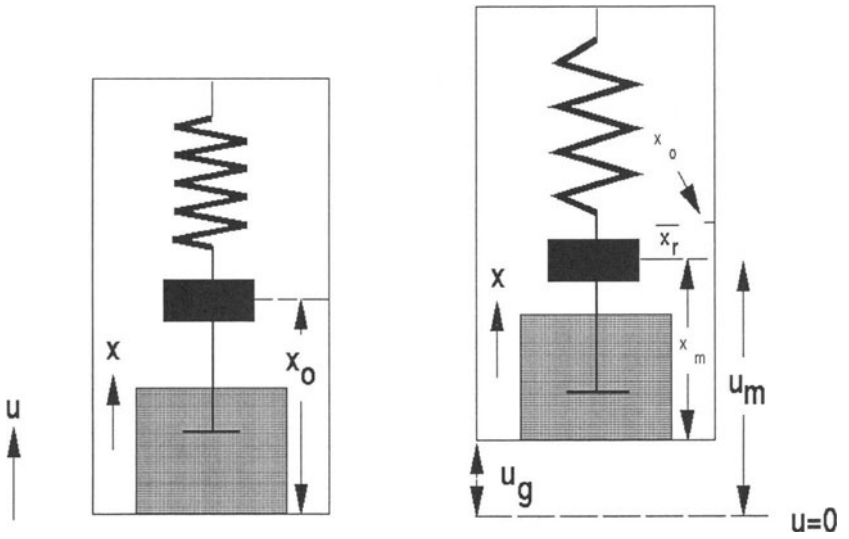



**Fig. 4.1** Model of a vertical pendulum seismometer. The inertial coordinate system is denoted **u**, while the **x** coordinate system is moving with the frame.

• *The spring* — A second force $f_{sp}$ is exerted by the spring, if the mass moves from its rest position $x_0$. $f_{sp}$ is proportional to the distance from the rest position $x_r = x_m - x_0$ measured within the reference frame $x$ which is attached to the seismometer frame. Since $f_{sp}$ is a restoring force, its sign is opposite to the sign of the displacement $x_r$.

$$f_{sp} = -kx_r(t)$$
$$k = \text{spring constant (strength)} \tag{4.2}$$

• *The dashpot* — A frictional force $f_f$ also acts on the mass. It is proportional to the velocity $\dot{x}_m(t)$ with which the mass moves relative to the dashpot. As with the effects of the spring, this quantity is best described in the reference frame $x$ which is attached to the frame of the seismometer.

As a dragging force, the frictional force is directed opposite to the direction of the velocity.

$$f_f = -D\dot{x}_m(t)$$

$$D = \text{friction coefficient} \tag{4.3}$$

In equilibrium, all these forces add up to zero:

$$-m\ddot{u}_m(t) - D\dot{x}_m(t) - kx_r(t) = 0 \tag{4.4}$$

From Fig. 4.1 we can see that $u_m(t) = u_g(t) + x_m(t)$, thus we can write

$$-m(\ddot{u}_g(t) + \ddot{x}_m(t)) - D\dot{x}_m(t) - kx_r(t) = 0 \tag{4.5}$$

Since $\dot{x}_m(t) = \dot{x}_r(t)$ and $\ddot{x}_m(t) = \ddot{x}_r(t)$, we can write (4.5) in terms of $x_r$, its derivatives and $u_g$:

$$m\ddot{x}_r(t) + D\dot{x}_r(t) + kx_r(t) = -m\ddot{u}_g(t) \tag{4.6}$$

If we divide by $m$, we obtain the equation of motion for the seismometer

$$\ddot{x}_r(t) + \frac{D}{m}\dot{x}_r(t) + \frac{k}{m}x_r(t) = -\ddot{u}_g(t) \tag{4.7}$$

Rewriting the constant terms we can write

$$\ddot{x}_r(t) + 2\varepsilon\dot{x}_r(t) + \omega_o^2 x_r(t) = -\ddot{u}_g(t) \tag{4.8}$$

where $\omega_o^2 = \frac{k}{m}$ and $2\varepsilon = \frac{D}{m} = 2h\omega_0$. $h = \frac{\varepsilon}{\omega_0}$ is called the damping constant of the seismometer.

From equation (4.8) we can already deduce some important characteristics of the seismometer:

— For rapid movements of the mass, the acceleration will be high compared to the velocity and the displacement. In this case $\ddot{x}_r$ will dominate the left hand side of (4.8) and the seismometer measures ground displacement.

— For slow movements of the mass (small acceleration and velocity), the $\ddot{x}_r$ and $\dot{x}_r$ become negligible and $x_r$ dominates the left hand side of (4.8). This corresponds to the seismometer measuring ground acceleration $\ddot{u}_g$.

To obtain the true motion of the ground for the general case, we must calculate a weighted sum of the relative movement between the moving mass and the seismometer frame ($x_r$) and its first and second time derivative ($\dot{x}_r$ and $\ddot{x}_r$, respectively). Therefore, we need to know the damping factor $\varepsilon$ or damping constant $h$, as well as the natural period of the seismometer $\omega_0$.

### 4.1 The solution for simple initial conditions

One way to determine the damping factor and the natural period of a seismometer is by releasing the seismometer mass from a known starting displacement at time t = 0 (release test). For an electrodynamic system this is normally done by applying a step in the current to the calibration coil. The corresponding initial conditions are:

$$\text{starting displacement: } x_r(0) = x_{r0}$$

$$\text{mass is at rest at time t = 0: } \dot{x}_r(0) = 0$$

$$\text{ground excitation is zero: } \ddot{u}_g(t) = 0$$

Equation (4.8) then becomes the homogeneous second order differential equation of the damped harmonic oscillator:

$$\ddot{x}_r(t) + 2\varepsilon\dot{x}_r(t) + \omega_0^2 x_r(t) = 0 \tag{4.9}$$

Making the classic trial solution: $x_r(t) = Ae^{\alpha t}$ with $\dot{x}_r(t) = \alpha A e^{\alpha t}$ and $\ddot{x}_r(t) = \alpha^2 A e^{\alpha t}$, we obtain

$$(\alpha^2 + 2\varepsilon\alpha + \omega_0^2)Ae^{\alpha t} = 0 \tag{4.10}$$

Since $e^{\alpha t} \neq 0$ for all $t$ we get

$$\alpha^2 + 2\varepsilon\alpha + \omega_o^2 = 0 \tag{4.11}$$

and the two solutions:

$$\begin{aligned} \alpha_1 &= -\varepsilon + \sqrt{\varepsilon^2 - \omega_0^2} \\ \alpha_2 &= -\varepsilon - \sqrt{\varepsilon^2 - \omega_0^2} \end{aligned} . \tag{4.12}$$

From the theory of linear differential equations we know that any linear combination of solutions is again a solution of the differential equation. Therefore, the general solution of equation (4.9) can be written as:

$$\begin{aligned} x_r(t) &= A_1 e^{\alpha_1 t} + A_2 e^{\alpha_2 t} = \\ &= A_1 e^{-(\varepsilon - \sqrt{\varepsilon^2 - \omega_o^2})t} + A_2 e^{-(\varepsilon + \sqrt{\varepsilon^2 - \omega_0^2})t} \end{aligned} \tag{4.13}$$

We can estimate the coefficients $A_1$ and $A_2$ from the initial conditions:

$$\begin{aligned} x_r(0) &= x_{r0} = A_1 + A_2 \\ \dot{x}_r(0) &= 0 = \alpha_1 A_1 + \alpha_2 A_2 \end{aligned} \tag{4.14}$$

We get

$$\begin{aligned} \alpha_1 A_1 &= -\alpha_2 A_2 = -\alpha_2(x_{r0} - A_1) = -x_{r0}\alpha_2 + \alpha_2 A_1 \\ A_1(\alpha_1 - \alpha_2) &= -x_{ro}\alpha_2 \end{aligned} \tag{4.15}$$

and the coefficients $A_1$ and $A_2$ are:

$$A_1 = x_{r0}\frac{\alpha_2}{\alpha_2 - \alpha_1} = x_{r0}\frac{\varepsilon + \sqrt{\varepsilon^2 - \omega_0^2}}{2\sqrt{\varepsilon^2 - \omega_0^2}}$$

$$A_2 = x_{r0}\frac{-\alpha_1}{\alpha_2 - \alpha_1} = x_{ro}\frac{-\varepsilon + \sqrt{\varepsilon^2 - \omega_0^2}}{2\sqrt{\varepsilon^2 - \omega_0^2}} \tag{4.16}$$

The solution of equation (4.9) finally becomes:

$$x_r(t) = \frac{x_{r0}}{2\sqrt{\varepsilon^2 - \omega_0^2}} e^{-\varepsilon t}(\varepsilon e^{\sqrt{\varepsilon^2 - \omega_0^2}t} + \sqrt{\varepsilon^2 - \omega_0^2} e^{\sqrt{\varepsilon^2 - \omega_0^2}t}$$

$$- \varepsilon e^{-\sqrt{\varepsilon^2 - \omega_0^2}t} + \sqrt{\varepsilon^2 - \omega_0^2} e^{-\sqrt{\varepsilon^2 - \omega_0^2}t})$$

$$= x_{ro} e^{-\varepsilon t}\left(\frac{\varepsilon}{\sqrt{\varepsilon^2 - \omega_0^2}}\left(\frac{e^{\sqrt{\varepsilon^2 - \omega_0^2}t} - e^{-\sqrt{\varepsilon^2 - \omega_0^2}t}}{2}\right) + \right.$$

$$\left.\left(\frac{e^{\sqrt{\varepsilon^2 - \omega_0^2}t} + e^{-\sqrt{\varepsilon^2 - \omega_0^2}t}}{2}\right)\right) \tag{4.17}$$

Depending on the values of $\varepsilon$ and $\omega_0$ we can distinguish three different cases:

### 4.1.1 UNDERDAMPED CASE

In this case $\omega_0 > \varepsilon$ and $\sqrt{\varepsilon^2 - \omega_0^2}$ becomes imaginary, namely $j\sqrt{\omega_0^2 - \varepsilon^2}$. Equation (4.17) becomes:

$$x_r(t) = x_{r0}e^{-\varepsilon t}\left(\frac{\varepsilon}{\sqrt{\omega_0^2 - \varepsilon^2}}\left(\frac{e^{j\sqrt{\omega_0^2 - \varepsilon^2}t} - e^{-j\sqrt{\omega_0^2 - \varepsilon^2}t}}{2j}\right)\right.$$

$$\left. + \left(\frac{e^{j\sqrt{\omega_0^2 - \varepsilon^2}t} + e^{-j\sqrt{\omega_o^2 - \varepsilon^2}t}}{2}\right)\right) \tag{4.18}$$

With Euler's formulas: $\cos y = \frac{e^{jy} + e^{-jy}}{2}$ and $\sin y = \frac{e^{jy} - e^{-jy}}{2j}$ we can rewrite equation (4.18) to:

$$x_r(t) = x_{r0}e^{-\varepsilon t}\left(\frac{\varepsilon}{\sqrt{\omega_0^2 - \varepsilon^2}}\sin\left(\sqrt{\omega_0^2 - \varepsilon^2}t\right) + \cos\left(\sqrt{\omega_0^2 - \varepsilon^2}t\right)\right) \tag{4.19}$$

With the definition $\sin\phi \equiv h = \frac{\varepsilon}{\omega_0}$ we obtain

$$\frac{\sin\phi}{\cos\phi} = \frac{\varepsilon/\omega_0}{\sqrt{1 - \varepsilon^2/\omega_0^2}} = \frac{\varepsilon}{\sqrt{\omega_0^2 - \varepsilon^2}} \tag{4.20}$$

and

$$x_r(t) = x_{r0}e^{-\varepsilon t}\left(\frac{\sin\phi}{\cos\phi}\sin\left(\sqrt{\omega_0^2 - \varepsilon^2}t\right) + \cos\left(\sqrt{\omega_0^2 - \varepsilon^2}t\right)\right)$$

$$= \frac{x_{ro}e^{-\varepsilon t}}{\cos\phi}\left(\sin\phi\sin\left(\sqrt{\omega_0^2 - \varepsilon^2}t\right) + \cos\phi\cos\left(\sqrt{\omega_0^2 - \varepsilon^2}t\right)\right) \tag{4.21}$$

With $\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$ and $\omega \equiv \sqrt{\omega_0^2 - \varepsilon^2}$ we get:

$$x_r(t) = \frac{x_{r0}}{\cos\phi} e^{-\varepsilon t}\cos(\sqrt{\omega_0^2 - \varepsilon^2}\,t - \phi)$$

$$= \frac{x_{r0}}{\cos\phi} e^{-\varepsilon t}\cos(\omega t - \phi)$$

$$\phi = arc\sin\left(\frac{\varepsilon}{\omega_0}\right) \tag{4.22}$$

In the underdamped case ($h < 1$), the seismometer oscillates with the period $T = \frac{2\pi}{\omega}$ which is always larger than the undamped natural period $T_0$:

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{\omega_0^2 - \varepsilon^2}} = \frac{2\pi}{\omega_0\sqrt{1 - \varepsilon^2/\omega_0^2}} = \frac{2\pi}{\omega_0}\frac{1}{\sqrt{1 - h^2}}$$

$$= \frac{T_0}{\sqrt{1 - h^2}} \tag{4.23}$$

### 4.1.2 OVERDAMPED CASE

For the overdamped case ($\omega_0 < \varepsilon$), the damping constant $h$ becomes greater than 1 and the solution of (4.13) becomes:

$$x_r(t) = A_1 e^{-(\varepsilon - \sqrt{\varepsilon^2 - \omega_0^2})t} + A_2 e^{-(\varepsilon + \sqrt{\varepsilon^2 - \omega_0^2})t}$$

$$= A_1 e^{-c_1 t} + A_2 e^{-c_2 t} \tag{4.24}$$

Since both $c_1$ and $c_2$ are real and positive, the solution will be an decaying exponential function. It will never oscillate.

### 4.1.3 CRITICALLY DAMPED CASE

For the critically damped case ($\omega_0 = \varepsilon$) $\omega_0 \to \varepsilon \Rightarrow \sin(\sqrt{\omega_0^2 - \varepsilon^2}t) \to \sqrt{\omega_0^2 - \varepsilon^2}t$ and $\cos(\sqrt{\omega_0^2 - \varepsilon^2}t) \to 1$. Hence (4.19) returns to:

$$x_r(t) = x_{r0}(\varepsilon t + 1)e^{-\varepsilon t} \tag{4.25}$$

### 4.1.4 COMPARISON

In Fig. 4.2, the output signals for an initial displacement of -1 units corresponding to three different values of the damping constant are shown

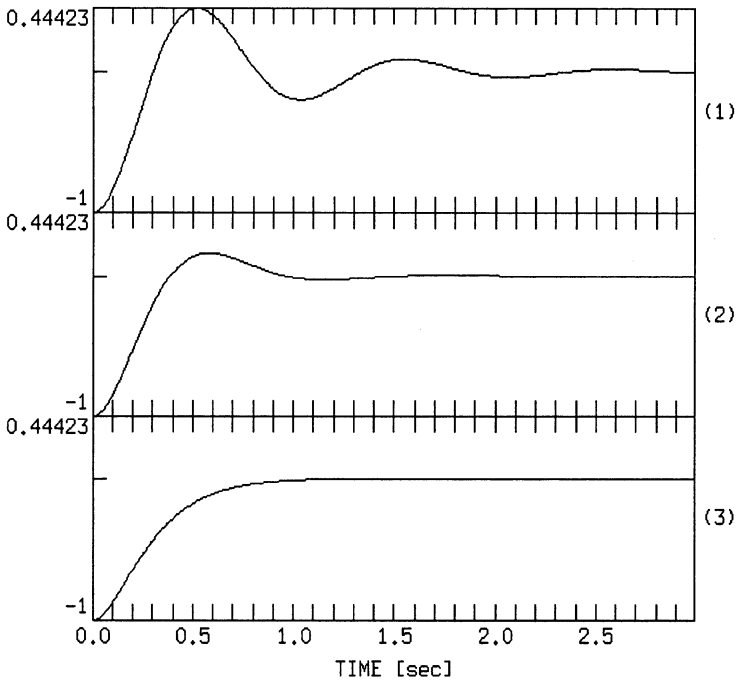



**Fig. 4.2** Dependence of the output signal of a displacement seismometer ($f_0$ = 1 Hz) on the damping constant $h$. From top to bottom $h$ changes from 0.25, 0.5, to 1.0. The initial displacement was assumed to be -1.

### 4.2 The determination of the damping constant

For the determination of the damping constant, we analyse the amplitude ratios of consecutive extreme values $a_k$ (maxima and minima) of the release test described above. Let us rewrite (4.19)

$$x_r(t) = x_{r0}e^{-\varepsilon t}(c_1 \sin(\sqrt{\omega_0^2 - \varepsilon^2}t) + c_2 \cos(\sqrt{\omega_0^2 - \varepsilon^2}t)) \tag{4.26}$$

Since the amplitude ratio of two consecutive maxima **or** two consecutive minima $(a_k, a_{k+2})$ is solely determined by the exponential term in equation (4.26), we can use this ratio to estimate the damping constant. We obtain:

$$\frac{a_k}{a_{k+2}} = \frac{e^{-\varepsilon t}}{e^{-\varepsilon(t+T)}} = \frac{e^{-\varepsilon t}}{e^{-\varepsilon t}e^{-\varepsilon T}} = e^{\varepsilon T} \tag{4.27}$$

$$ln\left(\frac{a_k}{a_{k+2}}\right) = \varepsilon T = \Lambda = \text{logarithmic decrement}$$

In terms of the amplitude ratio of consecutive extreme values $(a_k, a_{k+1})$ we obtain

$$\frac{a_k}{a_{k+1}} = \frac{e^{-\varepsilon t}}{e^{-\varepsilon(t+T/2)}} = e^{\varepsilon(T/2)} \tag{4.28}$$

$$\ln\left(\frac{a_k}{a_{k+1}}\right) = \frac{\varepsilon T}{2} = \frac{\Lambda}{2}$$

$$\Lambda = 2\ln\left(\frac{a_k}{a_{k+1}}\right) \tag{4.29}$$

Between the logarithmic decrement $\Lambda$ and the damping constant $h$, the following relationship exists:

$$\Lambda = \varepsilon T = \frac{\varepsilon T_0}{\sqrt{1-h^2}} = \frac{\varepsilon \frac{2\pi}{\omega_0}}{\sqrt{1-h^2}} = \frac{2\pi h}{\sqrt{1-h^2}} \tag{4.30}$$

$$h = \frac{\Lambda}{\sqrt{4\pi^2 + \Lambda^2}} \tag{4.31}$$

*Problem 4.1* Most seismometers operate on the principle of a moving coil within a magnetic field. Hence, they do not record the ground displacement but the ground velocity. Are equations (4.27) - (4.30) also valid for this kind of systems?

### 4.2.5 THE ELECTROMAGNETIC SENSOR

Most commonly, the pendulum motion $x_r(t)$ (cf. Fig. 4.1) is measured by an electrodynamic sensor (Fig. 4.3)

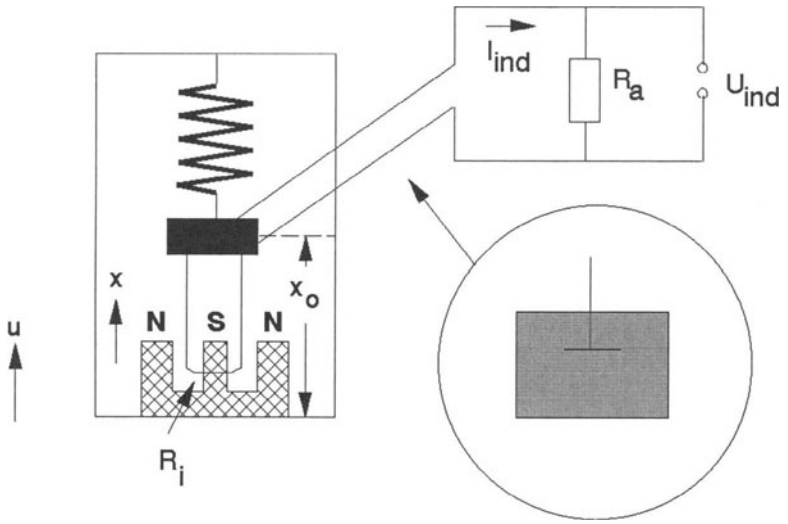



**Fig. 4.3** Schematic model of an electromagnetic sensor. The dashpot in figure 4.1 is replaced by a coil moving in a magnetic field.

In this case, the dashpot in Fig. 4.1 is replaced by a coil attached to the mass which moves through a permanent magnetic field. A voltage $U_{ind}$ is generated across the coil proportional to $\dot{x}_r(t)$, the velocity of the seismometer mass with respect to the seismometer frame. If the coil is shunted by a resistance $R_a$, the generated current $I_{ind}$ will be:

$$I_{ind} = \frac{U_{ind}}{R_a + R_i} \tag{4.32}$$

Here $R_i$ is the internal resistance of the damping circuit including the coil. The induced magnetic field will be oriented in a way to damp the motion producing the voltage $U_{ind}$. The damping factor $\varepsilon_c$ due to the coil is proportional to $\frac{1}{R_a + R_i}$.

$$\varepsilon_c \sim \frac{1}{R_a + R_i} \tag{4.33}$$

Taking into account the mechanical attenuation of the pendulum ($\varepsilon_0$) as well, we obtain ($h = \varepsilon/\omega_0$):

$$\varepsilon = h\omega_0 = \varepsilon_0 + b\frac{1}{R_a + R_i} \tag{4.34}$$

For the damping constant $h$, this becomes

$$h = h_0 + b'\frac{1}{R_a + R_i} \tag{4.35}$$

if $h_0$ describes the mechanical damping of the pendulum.

In addition to the elements shown in Fig. 4.3, electromagnetic seismometers often have a calibration coil to impose a predefined displacement onto the seismometer mass. Switching the calibration current on/off produces a step function in acceleration which is equivalent to releasing the seismometer mass from a known initial displacement. An example of a response to such a calibration signal is shown in Fig. 4.4.

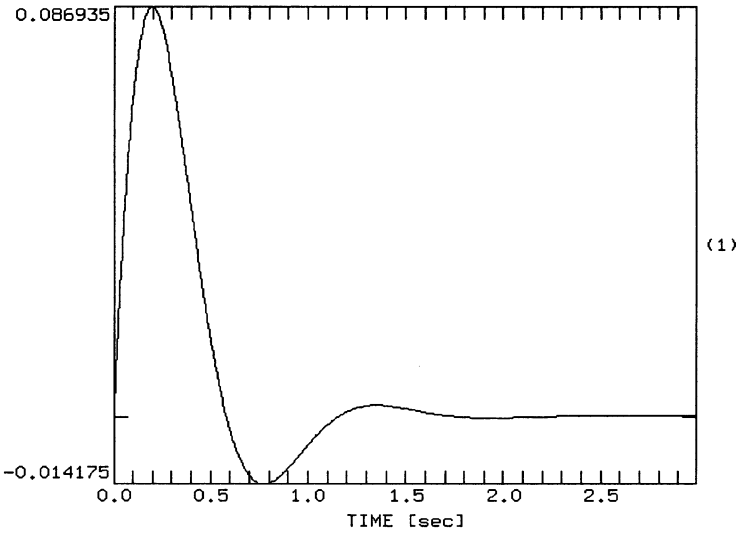



**Fig. 4.4** Seismometer calibration pulse (response of an electrodynamic seismometer to a step function in acceleration)

*Problem 4.2* The calibration signal shown in Fig. 4.4 is the response of a "velocity proportional sensor" to a step function in acceleration. Read the first two peak amplitude values $a_k$ and the damped period from Fig. 4.4. Determine the damping constant $h$ and the natural frequency $f_0$ of the system.

### 4.3 The frequency response function

The solution of the seismometer equation (4.8) under the initial conditions described in chapter *4.1* provided us with a tool to obtain the system's parameters. What we actually want to know, however, is the response of a seismometer to an arbitrary input signal. Since an arbitrary function can be described as a superposition of harmonics under very general conditions (Fourier series or Fourier integral), we will look in more detail at the solution of the seismometer equation if the input (ground motion) is a harmonic signal $u_g(t) = A_i e^{j\omega t}$. The ground acceleration corresponding to such a signal is $\ddot{u}_g(t) = -\omega^2 A_i e^{j\omega t}$ and equation (4.8) becomes:

$$\ddot{x}_r(t) + 2\varepsilon\dot{x}_r(t) + \omega_0^2 x_r(t) = \omega^2 A_i e^{j\omega t} \tag{4.36}$$

Again, we propose a solution of the form:

$$\begin{aligned} x_r(t) &= A_o e^{j\omega t} \\ \dot{x}_r(t) &= j\omega A_o e^{j\omega t} \\ \ddot{x}_r(t) &= -\omega^2 A_o e^{j\omega t} \end{aligned} \tag{4.37}$$

In general, $A_i$ and $A_o$ may be complex quantities. Inserting (4.37) into (4.36), we get:

$$-\omega^2 A_o + 2\varepsilon j\omega A_o + \omega_0^2 A_o = \omega^2 A_i . \tag{4.38}$$

Solving for $A_0/A_i$ we get

$$\frac{A_o}{A_i} = \frac{\omega^2}{\omega_0^2 - \omega^2 + j2\varepsilon\omega} = T(j\omega) \tag{4.39}$$

$T(j\omega)$ is the frequency response function of the seismometer. We can separate the real and imaginary parts:

$$T(j\omega) = \frac{\omega^2}{\omega_0^2 - \omega^2 + j2\varepsilon\omega} \cdot \frac{\omega_0^2 - \omega^2 - j2\varepsilon\omega}{\omega_0^2 - \omega^2 - j2\varepsilon\omega} \tag{4.40}$$

$$= \frac{\omega^2 \cdot (\omega_0^2 - \omega^2 - j2\varepsilon\omega)}{(\omega_0^2 - \omega^2)^2 - j^2 \cdot 4\varepsilon^2\omega^2} = \frac{\omega^2 \cdot (\omega_0^2 - \omega^2 - j2\varepsilon\omega)}{(\omega_0^2 - \omega^2)^2 + 4\varepsilon^2\omega^2} \tag{4.41}$$

$$= \frac{\omega^2 \cdot (\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + 4\varepsilon^2\omega^2} - j \cdot \frac{\omega^2 \cdot 2\varepsilon\omega}{(\omega_0^2 - \omega^2)^2 + 4\varepsilon^2\omega^2} \tag{4.42}$$

Or we can write $T(j\omega)$ in terms of amplitude (using (4.39) with $|A_o/A_i| = |A_o|/|A_i|$)

$$|T(j\omega)| = \frac{\omega^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\varepsilon^2\omega^2}} \tag{4.43}$$

and phase

$$\Phi(\omega) = \operatorname{atan}\left(\frac{Im}{Re}\right) = \operatorname{atan}\left(\frac{-2\varepsilon\omega}{\omega_0^2 - \omega^2}\right). \tag{4.44}$$

Thus,

$$\begin{aligned} T(j\omega) &= |T(j\omega)| \cdot e^{j\Phi(j\omega)} \\ &= \frac{\omega^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\varepsilon^2\omega^2}} e^{j\operatorname{atan}\left(\frac{-2\varepsilon\omega}{\omega_0^2 - \omega^2}\right)}. \end{aligned} \tag{4.45}$$

We finally rewrite equation (4.45) in terms of the damping constant $h = \varepsilon/\omega_0$, which we can measure:

$$\begin{aligned} |T(j\omega)| &= \frac{\omega^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4h^2\omega_0^2\omega^2}} \\ &= \frac{\omega^2/\omega_0^2}{\sqrt{(1 - \omega^2/\omega_0^2)^2 + 4h^2\omega^2/\omega_0^2}} \\ &= \frac{1}{\sqrt{(\omega_0^2/\omega^2 - 1)^2 + 4h^2\omega_0^2/\omega^2}} \end{aligned} \tag{4.46}$$

and

$$\Phi(\omega) = arc\tan\frac{-2h\omega_0\omega}{\omega_0^2 - \omega^2} = arc\tan\frac{-2h\omega/\omega_0}{1 - \omega^2/\omega_0^2} \tag{4.47}$$

Note that for an harmonic input signal with $\omega = \omega_0$, the phase shift is $\pi/2$ independent of the damping.

For an electrodynamic system, the output voltage is - within the passband of the instrument (see chapter 10) - proportional to ground velocity (instead of displacement). In addition, it depends on the generator constant G of the seismometer coil. Equation (4.46) then becomes

$$|T(j\omega)| = G\frac{\omega^3}{\sqrt{(\omega_0^2-\omega^2)^2+4h^2\omega_0^2\omega^2}}$$

$$= G\frac{\omega^3/\omega_0^2}{\sqrt{(1-\omega^2/\omega_0^2)^2+4h^2\omega^2/\omega_0^2}} \tag{4.48}$$

$$G \equiv \frac{\text{output voltage}}{\text{ground velocity}} \qquad \left[\frac{\text{V}}{\text{m/ sec}}\right] \tag{4.49}$$

*Problem 4.3* What is the theoretical relationship between the response of a "velocity proportional sensor" to a step function in acceleration as shown in Fig. 4.4 and the displacement impulse response?

### 4.3.6 THE TRANSFER FUNCTION

To obtain an expression for the transfer function, let us solve the seismometer equation (4.8) using the Laplace transform. The Laplace transform of equation (4-8) $(\ddot{x}_r(t) + 2\varepsilon\dot{x}_r(t) + \omega_0^2 x_r(t) = -\ddot{u}_g(t))$ is:

$$s^2X_r(s) + 2\varepsilon sX_r(s) + \omega_0^2X_r(s) = -s^2U_g(s) \tag{4.50}$$

and thus

$$(s^2 + 2\varepsilon s + \omega_0^2)X_r(s) = -s^2U_g(s)\,. \tag{4.51}$$

The transfer function is therefore

$$T(s) = \frac{X_r(s)}{U_g(s)} = \frac{-s^2}{s^2+2\varepsilon s+\omega_0^2} \tag{4.52}$$

Notice again, that the frequency response function equation (4.39) can be obtained from the transfer function equation (4.52) by replacing $s$ by $j\omega$.

Since a quadratic equation $x^2 + bx + c = 0$ has the roots $x_{1,2} = -b/2 \pm \sqrt{b^2/4 - c}$, we get for the pole positions $p_{1,2}$ (roots of the denominator polynomial in (4.52)):

$$\begin{aligned} p_{1,2} &= -\varepsilon \pm \sqrt{\varepsilon^2 - \omega_0^2} \\ &= -h\omega_0 \pm \omega_0\sqrt{h^2 - 1} \\ &= -(h \pm \sqrt{h^2 - 1})\omega_0 \end{aligned} \tag{4.53}$$

For the underdamped case ($h$ <1) the pole position becomes

$$p_{1,2} = -(h \pm j\sqrt{1 - h^2})\omega_0 \tag{4.54}$$

with the pole distance from the origin

$$|p_{1,2}| = |(h \pm j\sqrt{1 - h^2})| \cdot |\omega_0| = \sqrt{h^2 + (1 - h^2)} \cdot |\omega_0| = |\omega_0| \tag{4.55}$$

Therefore, the poles of an underdamped seismometer are located in the left half of the s-plane in a distance of $|\omega_0|$ from the origin. The quantity $h \cdot |\omega_0|$ gives the distance from the imaginary axis (Fig. 4.5).

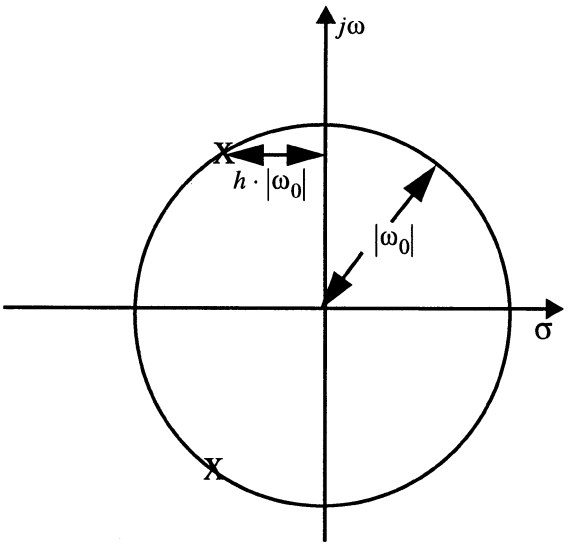



**Fig. 4.5** Geometrical interpretation of pole position, eigenfrequency $\omega_0$, and damping factor $h$ for a seismometer.

*Problem 4.4* Calculate the impulse response (and spectrum) for a displacement seismometer with an eigenfrequency of 1 Hz and damping factors of $h = 0.25, 0.5, 0.62$, respectively. How do the locations of the poles change for the different damping constants?

*Problem 4.5* How do we have to change the pole and zero distribution if we want to change the seismometer from Problem 4.4 into an electrodynamic system recording ground velocity?

CHAPTER 5

# The sampling process

Digital seismograms are sequences of numbers which in general have been obtained from the continuous output voltage of a seismic sensor by the procedures of sampling and analog to digital conversion. We will now consider some of the characteristic properties of these processes. We will simulate the sampling of analog data and its reconstruction from sampled values. We will find that an analog signal can only be reconstructed from its sampled values, if the frequency content of the signal to be sampled contains no energy at and above half of the sampling frequency (sampling theorem). We will investigate what happens if we deliberately violate this rule (aliasing effect).

## 5.1 The sampling of analog data

When we use a computer program to model continuous phenomena - like simulating the output voltage of a seismometer for certain boundary conditions - we usually do not think about the underlying process. We just do it and assume the results are meaningful. In terms of system theory, however, we made an important transition: From a continuous system to a discrete system. This step entails some rules we can not violate without jeopardizing the results.

The same transition from a continuous system to a discrete system takes place when we acquire data in digital form. In this transition there are actually two different steps:

• *Sampling or discretization* — Taking discrete samples of a continuous data stream. The data may still be in analog representation after the sampling process.

• *Analog to digital conversion (quantization and coding)* — For voltage signals, this steps normally occurs in an electronic device which is called ADC, 'analog to digital converter'. After having gone through this step, which we will treat in chapter 6, the data are digital and discrete.

In Fig. 5.1 the principle of the discretization process is displayed schematically. A continuous signal is sampled at discrete times indicated by the positions of the vertical arrows. The amplitudes of the arrows correspond to the signal amplitudes at the sampling

times. We can also view the discretization process as modulating a sequence of delta impulses which are separated in time by $T$. In the present context, only constant sampling intervals $T$ (equidistant discretization) are relevant. Its reciprocal, $1/T = f_{dig}$, is called the *sampling frequency* or the *digitization frequency.*

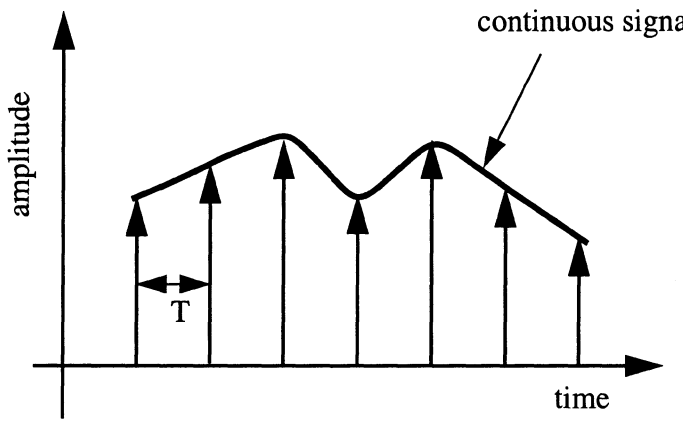



**Fig. 5.1** Sketch of the discretization (sampling) process. The vertical arrows show the locations and the values of the samples. $T$ denotes the sampling interval.

Below, some of the effects of sampling of 'continuous' signals are demonstrated using the *discretization tool* within DST. Of course, all the data traces in DST are already in digital form that is in form of a sequence of numbers. We are only approximating a continuous signal by one which has been sampled at a sampling frequency much higher than the one at which we want to investigate. Let us assume for now without proof that this is a valid approximation. From the information in this chapter it should become clear under which conditions this is permissible. Also, let us assume without proof that we can reverse the sampling process and recover an approximation of the original continuous signal. In DST, a procedure called *Whittaker reconstruction* is used for this task. For the details of this process, which we do not need to be concerned about for the purpose of the following argument, see for example Stearns (1975).

Fig. 5.2 displays a sinusoidal input signal with a signal frequency of 1 Hz which we will use for simulating the discretization process with DST.

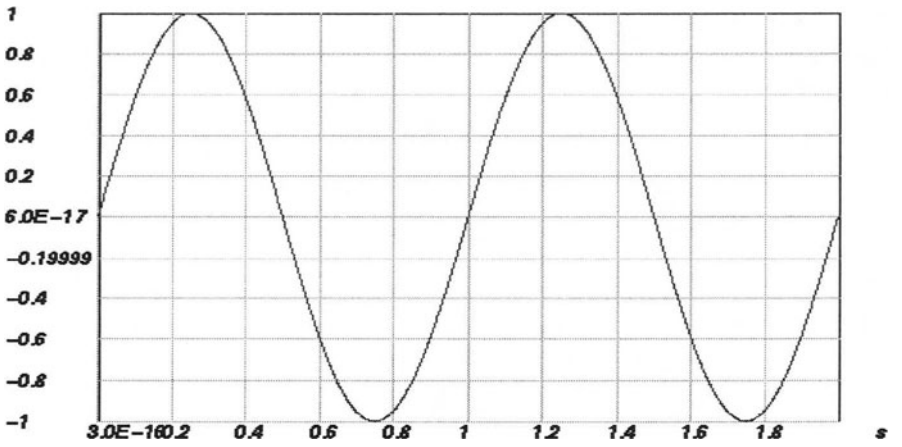



**Fig. 5.2** Input signal for the simulation of the discretization process. The signal frequency is 1 Hz.

During discretization, the 'continuous' input signals are 're-sampled' at the locations defined by the discretization frequency (Fig. 5.3)[1].

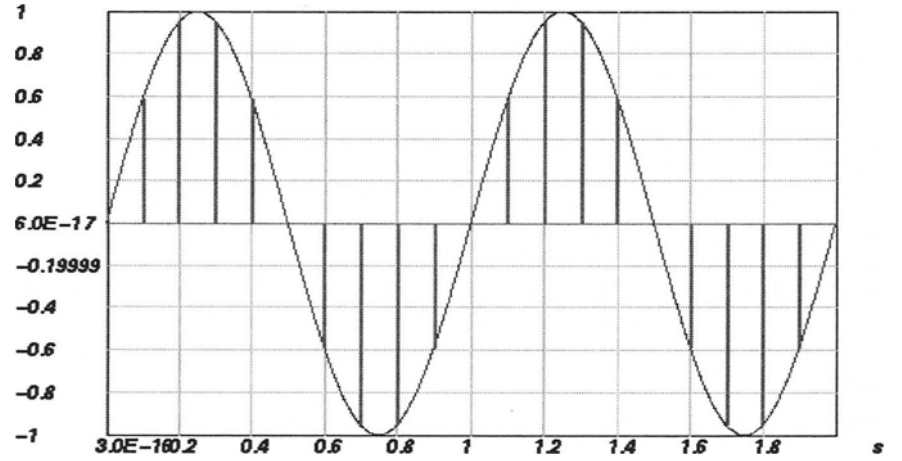



**Fig. 5.3** Discretizing the data trace of Fig. 5.2 using a discretization frequency of 10 Hz. The vertical bars show the locations and the values of the function at the sampled times.

1. It should be noted, that within DST the discretization process is simulated by zeroing out all but the 'discretized' samples which are those samples closest to the multiples of the 'discretization interval'. The internal sampling frequency is not actually altered during this process.

The times at which the vertical bars appear in Fig. 5.3 correspond to the times at which samples are taken during the discretization process for the given discretization frequency (here 10 Hz). The amplitudes of these bars corresponds to the trace amplitudes at those times.

Fig. 5.4 shows the result of discretizing the trace in Fig. 5.2 using a 'discretization frequency' of 10 Hz and reconstructing them again into a 'continuous' representation using a Whittaker reconstruction.

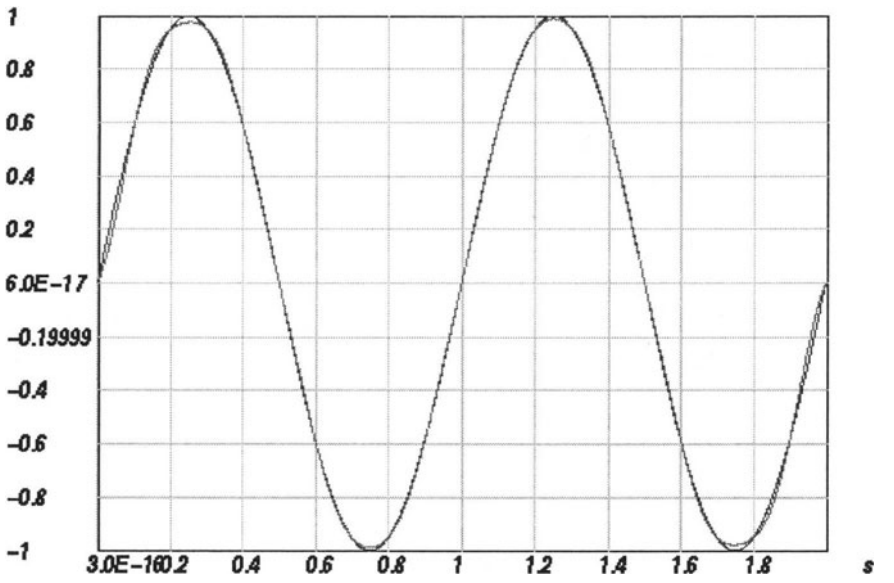



**Fig. 5.4** Original and reconstructed trace of Fig. 5.2 (after discretizing all of them with 10 Hz prior to reconstruction).

You can see from comparison of Fig. 5.4 and Fig. 5.2 that except for minor numerical distortions the 1 Hz sinusoidal signal is reconstructed correctly. However, using the same discretization frequency of 10 Hz for a sinusoidal signal with a signal frequency of 9 Hz yields a completely different result (Fig. 5.5).

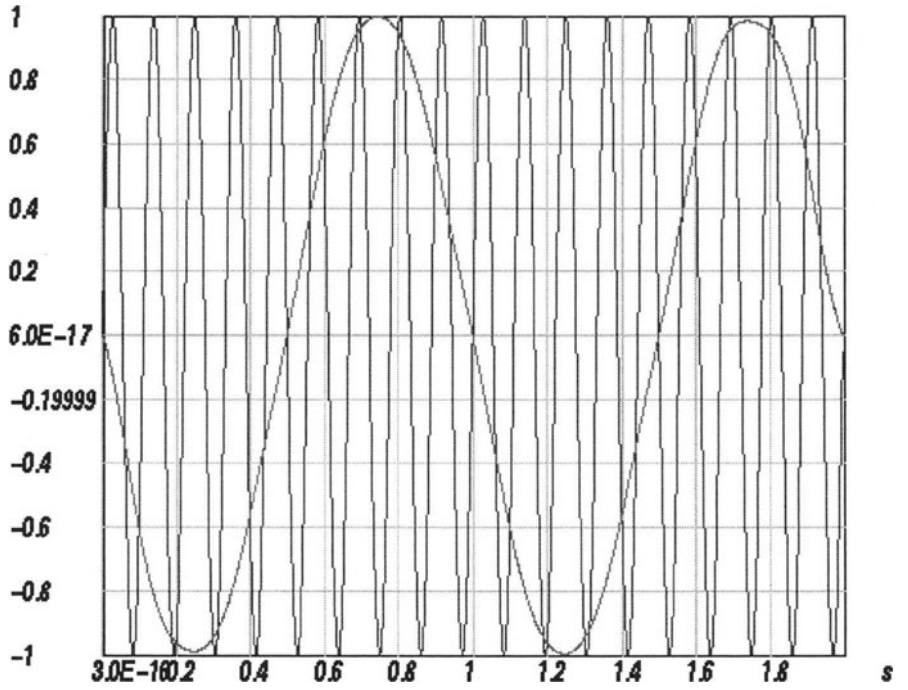



**Fig. 5.5** Original and reconstructed sinusoidal signal with a signal frequency of 9 Hz (discretization frequency 10 Hz).

The reconstructed signal is completely different from the original one. It shows two full oscillations in 2 seconds, hence the signal frequency of the reconstructed signal is 1 Hz instead of the 9 Hz of the original input signal. This is called *alias-effect* and the wrong frequency of the output signal is called alias-frequency.

Qualitatively, the alias effect shown in Fig. 5.5 can be understood from comparing the discretized signal (vertical bars) with the original input trace (solid line) as shown in Fig. 5.6. Obviously, the number of samples taken for each oscillation of the input trace is insufficient to reconstruct the original trace properly.

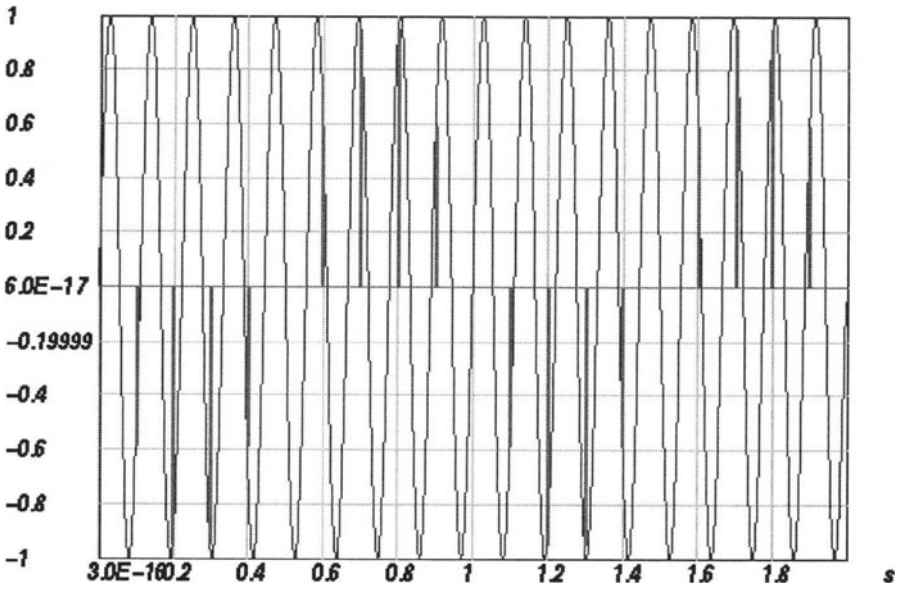



**Fig. 5.6** Discretizing a sinusoidal signal with a signal frequency a discretization frequency of 10 Hz. The vertical bars show the locations and the values of the function at the sampled times

The dominant frequency of the reconstructed signal (alias frequency) and the signal frequency of the input signal are related via the discretization frequency in a very systmatic way. This is treated in the following problems.

*Problem 5.1* What is the general relationship between the discretization frequency, the signal frequency of a sinusoidal input signal ("input frequency") and the dominant frequency of the reconstructed signal ("output frequency")? Use DST to generate sinusoidal signals for an internal sampling frequency of 1024 Hz, a window length of 2048 points, and signal frequencies from 1 - 20 Hz in steps of 1 Hz. Discretize and reconstruct each signal using a discretization frequency of 10 Hz and note the dominant frequency and the maximum amplitude of the reconstructed signal. You can determine the dominant frequency of a signal in Hz easily by measuring the dominant signal period on the DST screen in seconds and taking the reciprocal value. From the table of input frequencies, output frequencies and output amplitudes, try to infer the rule for calculating the output frequency for a given signal frequency and a given digitization frequency. Hint: The so called Nyquist frequency (half of the discretization frequency) is also referred to as the *folding frequency.* Think of the "frequency band" as a foldable band which is folded at multiples of the Nyquist frequency. Mark the corresponding pairs of (input frequency, alias frequency) on this band. It may help to actually cut out a paper band and folding it.

*Problem 5.2* What is the highest frequency which can be reconstructed correctly using a 'discretization frequency' of 10 Hz?

*Problem 5.3* What would be the alias frequency for an input signal of 18.5 Hz and a discretization frequency of 10 Hz?

Once a signal has been discretized at a discretization frequency too low for the actual frequency content, the signal can not be uniquely reconstructed. This is described by the *sampling theorem* given below.

> • *Sampling theorem* — For a continuous time signal to be uniquely represented by samples taken at a sampling frequency of $f_{dig}$, (every $1/f_{dig}$ time interval), no energy must be present in the signal at and above the frequency $f_{dig}/2$. $f_{dig}/2$ is commonly called the *Nyquist frequency* (e.g. Mitra and Kaiser, 1993). Signal components with energy above the Nyquist frequency will be mapped by the sampling process onto the so called alias frequencies within the frequency band of 0 to Nyquist frequency. This effect is called the *alias* effect.

An example of the alias effect can be seen in old western movies. If the wheels of a stagecoach seem to turn backwards, then the sampling of the images was to slow to catch the movement of the wheels uniquely.

In order to reconstruct a continuous signal from its sampled values, analog signals must be filtered so that all the unwanted frequency components are removed (**anti-alias filter**). The reconstruction of a continuous signal from its sampled values (e. g. the Whittaker reconstruction) is also done by lowpass filtering. The Whittaker reconstruction simply removes all frequencies above the Nyquist frequency.

It is important to stress that aliasing introduces frequencies into the undersampled signal which were not present in the original continuous signal. Consequently, sampling a signal at a sampling rate that is too low does not mean that the higher-frequency information is just lost, but means that the information in the resulting sampled signal is **wrong**. Care must be taken to choose the proper sampling rate in function of the frequency content of the continuous signal or by lowpass filtering the signal **before sampling it.** Once the signal has been sampled, there is no general way to determine whether aliasing might have occurred, and thus there is no way to correct for aliasing problems after the sampling process. The sampling process is also treated more theoretically in chapter 7.

CHAPTER 6

# Analog-to-digital conversion

In this chapter we will look at the properties of analog-to-digital converters (ADC) and the limitations they introduce into sampled data. We will discuss the relationship between the resolution and the dynamic range of ADCs and end this chapter by simulating techniques for improving the dynamic range of ADCs by gain ranging and oversampling.

## 6.1 The principle of analog to digital conversion

We have seen that the temporal characteristics of the sampling process imposes limits on the frequency content of the data to be sampled. We simulated the sampling process in DST, by selecting a subset of the original input sequence with amplitudes identical to the original values at the sampled times. When we convert a continuous analog input signal to a digital sequence of samples by using electronic analog-to-digital converters on voltage signals, this is no longer true. Through the process of analog-to-digital conversion, which involves the two processes of quantization and coding, the 'analog amplitudes' become a sequence of numbers (Fig. 6.1).[1]

1. Following the notation of Oppenheim and Schafer (1989), [] are used to enclose the independent variable of sequences. In contrast, () are used to enclose the independent variables of continuous-variable functions.

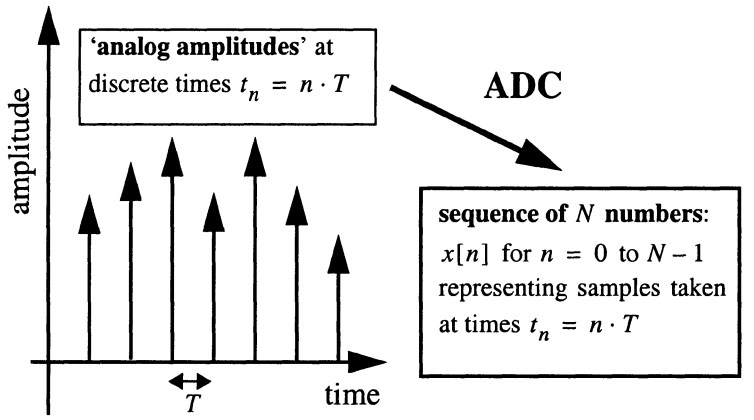



**Fig. 6.1** Sketch of the analog to digital conversion (ADC) process. The vertical arrows show the locations and the values of the samples. Through the ADC the amplitude values are converted into a sequence of numbers which represent the samples taken at times $t_n = n \cdot T$

To understand the sources of errors involved, we will look at the working principles of a simple type of analog to digital converter, the single slope ADC (Fig. 6.2)

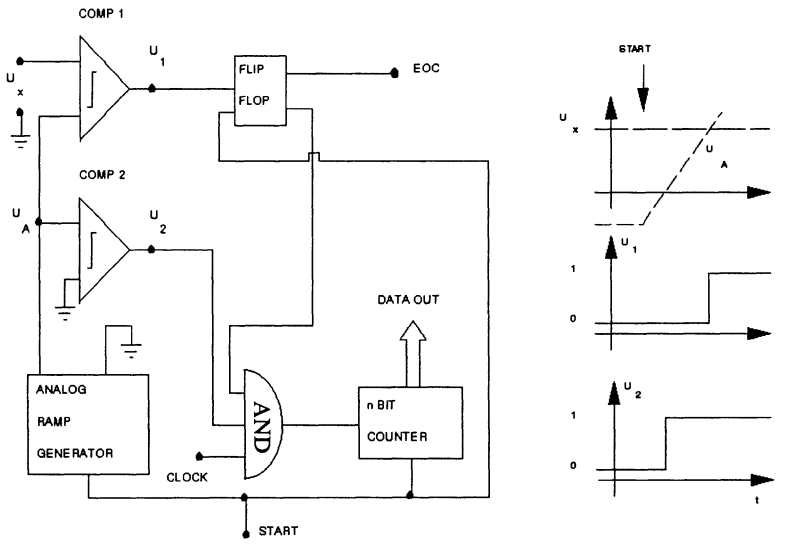



**Fig. 6.2** Principle of a single slope analog to digital converter (after Jaeger, 1982).

Assume that an input voltage $U_x$ is to be converted into a discrete, digital value. Let voltage $U_x$ remain constant during the time needed for the conversion. For this example we also assume (without loss of generality) $U_x$ to be positive. A start signal initiates the actual conversion and starts an analog ramp generator which produces a voltage $U_A$ which linearly increases with time (cf. Fig. 6.2 uppermost right panel). The voltage $U_A$ is checked by a comparator (COMP 2) to determine whether it is equal to or larger then 0 V. Once it reaches 0 V, output line $U_2$ of COMP 2, which is connected to a logic element (AND), goes high. The line between this logic element and the *n bit counter* will go high if all the input lines are high. In addition to the comparator 2, the remaining input lines of the logic element are connected to a clock (which goes high every clock cycle) and a flip flop. A flip flop is a bistable oscillator which can only have two different states (HIGH or LOW). Since the flip flop is assumed to start out in a high state, once the voltage $U_2$ is high, the counter will receive a voltage pulse every clock cycle. All the counter does is counting the pulses it receives. Since the counter uses the binary number code, the value of each bit of the counter represents a digit $d_i$ of a number expressed as sum of powers of 2 (Fig. 6.3). The maximum number which can be represented by 3 bits for example is $1 \cdot 2^0 + 1 \cdot 2^1 + 1 \cdot 2^2 = 2^3 - 1 = 7_{(10)}$ while the minimum number is 0. Likewise, if the counter has $n$ bits at its disposal to store the result of its counting process it can vary between 0 and $2^n - 1$.

Once the counting process has started, the voltage $U_A$ continues to increase. At a second comparator (COMP 1), it is continuously compared to the input voltage $U_x$. When $U_A \geq U_x$, COMP1 sets $U_1$ high and this toggles the flip flop off. This in turn prohibits the logic element from sending a signal to the counter. In other words, conversion stops.

**decimal system**

$$x_{(10)} = \sum_i d_i^{(10)} \cdot 10^i$$

e.g. $512_{(10)} = 2 \cdot 10^0 + 1 \cdot 10^1 + 5 \cdot 10^2$

least significant digit

most significant digit

conventional notation:

digits are written from most significant to least significant:

512

**binary system**

$$x_{(2)} = \sum_i d_i^{(2)} \cdot 2^i$$

e.g.

$13_{(10)} = 1 \cdot 2^0 + 0 \cdot 2^1 + 1 \cdot 2^2 + 1 \cdot 2^3$

least significant bit value

most significant bit value

conventional notation:

bit values are written from most significant (MSB) to least significant (LSB):

1101

**Fig. 6.3** Number representation in the decimal and binary number codes.

The input voltage $U_x$ is converted into a digital value by counting the time it takes for the ramp generator to produce a voltage as high as the input voltage. If the counter has $n$ bits for storing the result of its counting, the ADC is called a $n$-bit ADC, which means it has $2^n$ output states. Input voltages ranging from 0 to the full scale range voltage of the ADC are mapped onto discrete values between values between 0 and $2^n$ -1. For a 3 bit ADC and a full scale input voltage of 10 V, this situation is shown in Fig. 6.4.

The voltage corresponding to a change of the least significant bit of the counter is the smallest input voltage change that can cause a change of the output value of the ADC. It is called the quantum or simply least significant bit value (LSB value). For a $n$-bit ADC, it is given by

$$Q = \text{LSB value} = \frac{\text{FULL SCALE VOLTAGE}}{2^n} \tag{6.1}$$

Since the value of the quantum which controls the resolution of the ADC depends directly on the number of bits of the ADC, a $n$-bit ADC is said to have $n$ bits of *resolution*. For the linearly increasing input signal in Fig. 6.4, the probability density function of the error signal is flat between $\pm Q/2$. Consequently, as a uniformly distributed white-noise sequence it has a variance of $Q^2/12$ (Oppenheim and Schafer, 1989, p. 120). As can be seen from Fig. 6.4, the relative error of the ADC depends on the level of the input signal. Therefore, the assumption that the quantization error is uniformly distributed between $\pm Q/2$ does only hold in general when $Q$ is small and the signal traverses several quantization levels between successive samples (Proakis and Manolakis,1992).

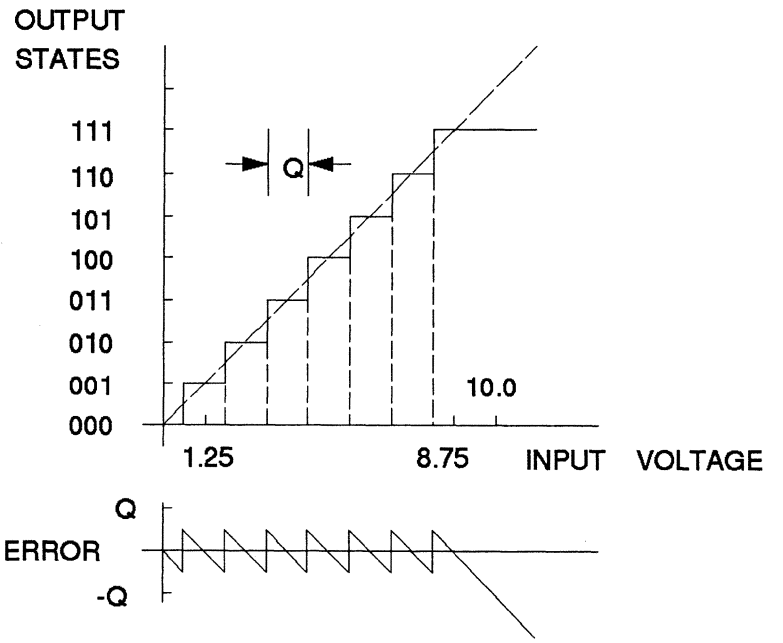



**Fig. 6.4** Mapping of input voltage to output states for a 3 bit ADC. The lower panel shows the error signal. Q corresponds to the quantum or least significant bit value (LSB value).

An additional measure of the quality of an ADC is its *dynamic range*. In general, dynamic range is defined as the ratio between the largest ($A_{max}$) and smallest ($A_{min}$) amplitude which can be measured from an instrument. In the context of this definition,

amplitudes $A_{max}$ and $A_{min}$ are sometimes defined in the frequency and sometimes in the time domain. In either case, dynamic range is commonly expressed in decibels ($20 \cdot \log_{10}$(amplitude ratio)).

$$D = 20\log_{10}(A_{max}/A_{min}) \qquad [\text{dB}] \tag{6.2}$$

Besides this general definition of dynamic range which can be used for any analog signal, for discrete, digitized signals dynamic range can also be defined by the range of numbers which can be represented with a given number of bits.

Consequently, a $n$-bit ADC which can represent a maximum number of $2^n - 1$ numbers can be said to have a dynamic range of

$$D = 20\log_{10}(2^n - 1) \approx n \cdot 20\log_{10}(2) \approx n \cdot 6 \qquad [\text{dB}], \tag{6.3}$$

with 1 bit corresponding to approximately 6 dB.

Thus, the ADC in Fig. 6.4 is said to have a resolution of 3 bits and a dynamic range of 18 dB. Fig. 6.5 demonstrates what happens if an input signal has larger amplitudes than the ADC can represent. The upper trace in Fig. 6.5 shows a decaying sinusoid with a maximum amplitude of 839.46 (arbitrary units). Below is the output signal for an ADC with a resolution of 12 bit and a LSB of 0.1. Notice that the output trace is saturated at an amplitude of $0.1 \cdot (2^{12} - 1) = 0.1 \cdot 2047 = 204.7$.

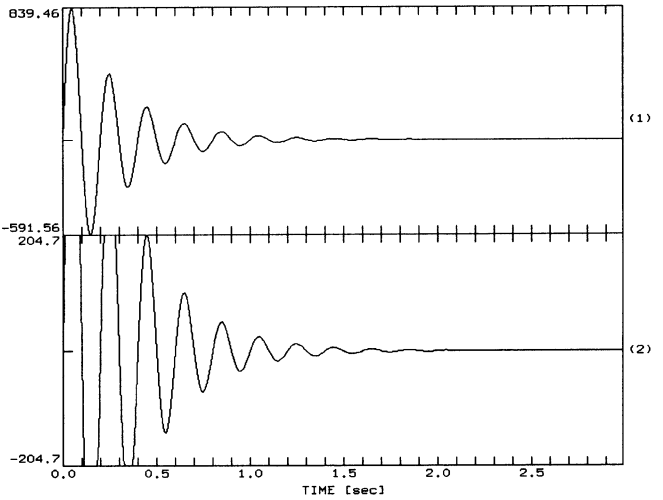



**Fig. 6.5** The saturation effect for an ADC with insufficient dynamic range. For an ADC with a resolution of 12 bit and a LSB of 0.1 the output trace is saturated at an amplitude of $0.1 \cdot (2^{12} - 1) = 0.1 \cdot 2047 = 204.7$.

A simple work-around for the saturation (clipping) problem would be to attenuate the input channel before it is fed into the ADC as shown in Fig. 6.6

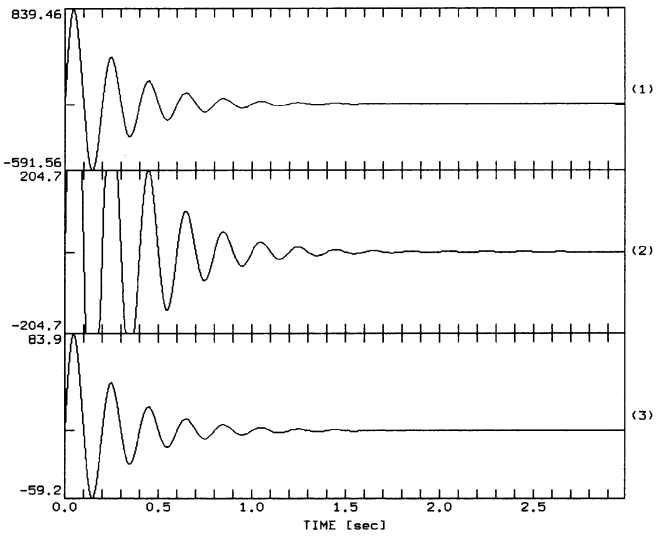



**Fig. 6.6** Removing the saturation effect by attenuating the input signal before analog to digital conversion using an attenuation factor of 10 (cf. channel 3 and 2).

In Fig. 6.6, the ADC input signal (top trace) is attenuated by a factor of 10 before being fed into the ADC with the same parameters as being used for Fig. 6.5. As we see in the lowermost trace, the dynamic range is now sufficient to convert the complete waveform without saturation. However, we run into a different problem using this simple minded approach. Fig. 6.7 shows the later portion of the original signal (top) and the attenuated resampled signal (bottom).

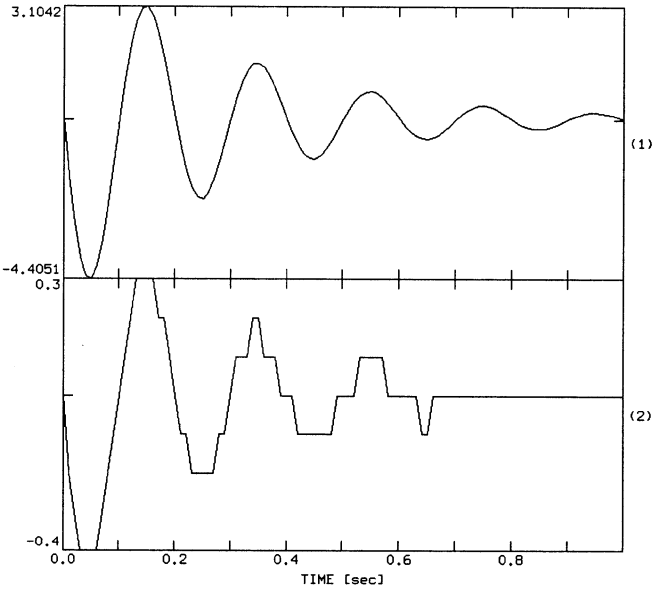



**Fig. 6.7** Lack of resolution. The value of $Q$ determines the smallest amplitude difference which can be resolved during analog to digital conversion. For the attenuated trace (channel 2) the amplitude differences within the selected time window become too small for the given $Q$ to be resolved properly.

For this segment of data, most of the changes in the attenuated signal amplitude are below what the ADC can resolve (LSB of 0.1). Hence, we have traded the maximum value we can record (which helped with the big amplitudes) with a lack of 'resolution' for the small amplitudes.

*Problem 6.1* The Wood-Anderson magnitude is defined as $M_{WA} = \log_{10}(A) - \log_{10}(A_0)$ with $A$ being the amplitude [in mm] measured on a Wood-Anderson displacement instrument and $-\log_{10}(A_0)$ being the distance correction which is exactly 3 for 100 km. What is the required dynamic range of a digital Wood-Anderson-equivalent instrument to record both a magnitude 0 and a magnitude 6 earthquake in 100 km distance "on scale"? Hint: Consider the theoretical Wood-Anderson trace amplitudes for both earthquakes in 100 km distance.

## 6.2 Increasing dynamic range and resolution

The dynamic range needed for recording a reasonable magnitude range on scale still touches the limits of currently available, reasonably priced ADCs. Nonetheless, people have found ways to improve both the dynamic range as well as the resolution of recording using currently available technology. We will now discuss and simulate two different methods: *gain ranging* and *oversampling*. The first technique trades dynamic range for resolution while the second technique increases the resolution by reducing the influence of quantization noise and consequently increasing the dynamic range.

### 6.2.1 GAIN RANGING

Fig. 6.8 sketches the structure of a gain ranging ADC. The key elements are a regular n-bit ADC, a programmable gain amplifier (PGA), and a control logic. As long as the analog input signal is small in comparison with the full-scale range of the ADC, the gain ranging ADC works like a plain ADC. However, once the input signal reaches a certain level ('switch up fraction', e.g 0.25 of the FULL SCALE RANGE of the ADC) the control logic will cause the PGA to decrease the gain of the pre-amplifier. As a consequence, the signal will again fit into the full scale range of the ADC easily. If the signal amplitude increases further, the gain of the pre-amplifier will be switched down again. This can continue until the lowest gain is reached.

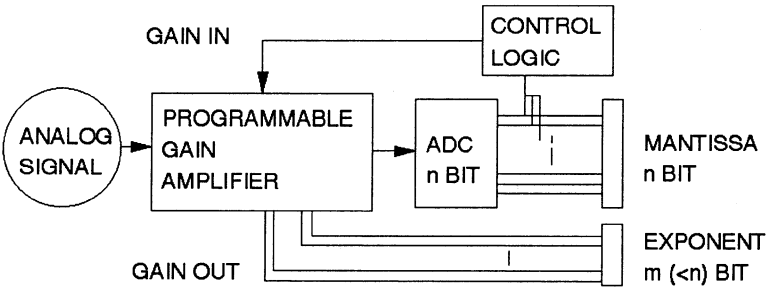



**Fig. 6.8** The principle of a gain ranging ADC.

On the other hand, if the signal amplitude has been below a certain threshold for a certain number of cycles, the control logic will cause the PGA to raise the gain again.

In a gain ranging ADC, both the output values of the plain ADC (mantissa) and the status of the PGA (gain) are recorded. If we use $m$ bits to record the gain status, we can record $2^m$ gain states from 0 to $2^m - 1$. Since the PGA switches gain in powers of 2, the maximum gain factor is $2^{(2^m - 1)}$. The dynamic range of a gain ranging ADC with m bits for

the gain and n bits for the mantissa can therefore be defined as:

$$\begin{aligned} D_{gr} &= 20\log_{10}((2^n - 1) \cdot 2^{(2^m - 1)}) \quad [\text{dB}] \\ &\approx 20\log_{10}(2^n \cdot 2^{2^m}) \quad [\text{dB}] \\ &= 20\log_{10}(2^{n+2^m}) \quad [\text{dB}] \end{aligned} \tag{6.4}$$

For a given number of total bits, the increase in dynamic range is accompanied by a gain dependent loss of resolution. In Fig. 6.9, the data trace in Fig. 6.5 is resampled using a gain ranging ADC with 8 bits of resolution for the mantissa and 4 bits for the exponent (negative gain). From the 16 ($2^4$) possible states of the programmable gain amplifier (PGA), however, only 11 are used in this example (in real gain ranging ADCs, some bits of the exponent are sometimes used as error flag). Visually we do not see a difference between the result of the gain ranging ADC (bottom trace) and the original input trace (top trace) even in the 'low amplitude' range of the signal. For comparison, the second and the third trace in Fig. 6.9 show the corresponding data window for trace 2 in Fig. 6.5 (insufficient dynamic range) and trace 2 in Fig. 6.7 (insufficient resolution), respectively. We can see that the result of the gain ranging ADC is visually indistinguishable from the input trace. The same is true for the high amplitude region.

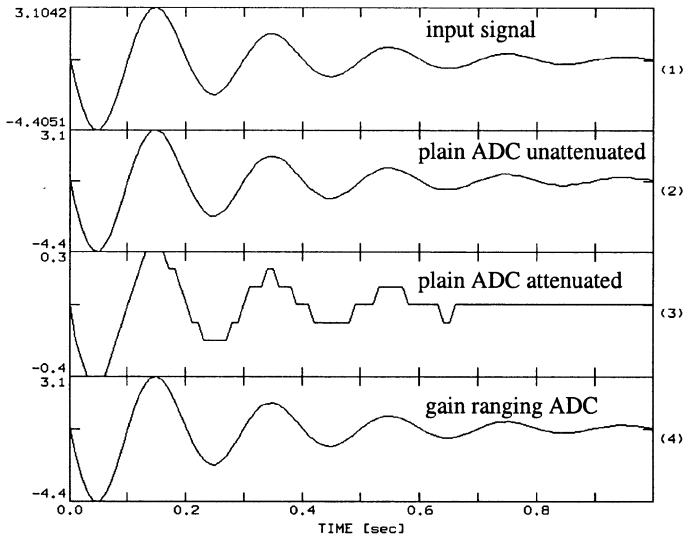



**Fig. 6.9** Comparing the resolution of plain- and gain ranging ADC for a 'low amplitude' time window. Notice that the second trace suffers from saturation for the 'high amplitude' time window (cf. Fig. 6.5).

As a final example for the action of gain ranging ADCs, Fig. 6.10 shows from top to bottom the input trace, the quantized values, the mantissa, the negative gain, and the error signal for a half wave of a cosine signal with a maximum amplitude of 2000. The resolution was chosen to be 12 bit, the LSB = 0.1, and the number of possible gain states as 11.

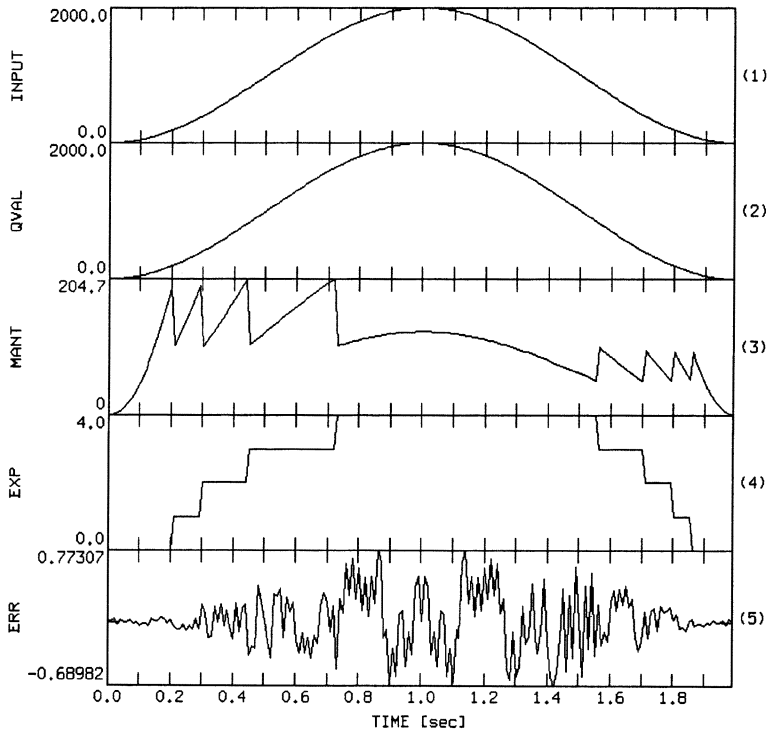



**Fig. 6.10** Input signal, quantized value, mantissa, exponent (gain reduction states), and error signal for a gain ranging ADC[1].

The quantized value in channel 2 equals (channel 3) x $2^{(\text{channel 4})}$. The error signal in channel 5 is the difference between the input signal and the quantized values. As a consequence of the trade-off between dynamic range and resolution for gain ranging ADCs, the absolute value of the error signal becomes dependent on the gain state.

*Problem 6.2* What are the differences in dynamic range (in dB) between a plain 16 bit ADC and a gain ranging ADC with 8 bit for the gain and 8 bit for the mantissa?

1. In the present example, the logic for the gain switching is different when the signal amplitude is increasing or decreasing. This is also the case in some real ADCs.

### 6.2.2 OVERSAMPLING

A set of techniques which has become very popular recently is based on sampling the seismic signals at much higher sampling rates than finally desired (oversampling). The sampled signals are subsequently digitally lowpass filtered and downsampled (decimated) to the desired sampling rate. In the following it is discussed how this increases the resolution of analog to digital conversion by reducing the level of quantization noise within the final frequency band.

Oversampling in its simplest form works on the assumption made earlier that the variance of the quantization noise is, under certain general conditions, independent of the sampling rate (Oppenheim and Schafer, 1989; Proakis and Manolakis, 1992). In this case the quantization noise has a probability density function which is flat between $\pm Q/2$ (cf. Fig. 6.2, Q = quantum of the ADC). Therefore, the variance of the quantization noise is $Q^2/12$ independent of the sampling rate. For $N$ samples of quantization noise $\varepsilon[n]$ the sample variance $s^2$ becomes

$$s^2 \equiv \frac{\sum_{n=1}^{N} |\varepsilon[n]|^2}{N} = \frac{Q^2}{12} = const \tag{6.5}$$

On the other hand, the variance $s^2$ corresponds to the total signal power of the error signal within a time window of length $T$ (Press et al., 1992)

$$s^2 \equiv \frac{\sum_{n=1}^{N} |\varepsilon[n]|^2}{N} \approx \frac{1}{T} \cdot \int_0^T |\varepsilon(t)|^2 dt \tag{6.6}$$

Before we can exploit this to increase the resolution of the ADC, we informally introduce the discrete Fourier transform (DFT) for a sampled signal $x[n]$:

$$\tilde{X}[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N} \tag{6.7}$$

Here, $x[n]$ is defined for $n = 0$ to $N-1$. The discrete Fourier transform will be introduced more formally and placed in context in the following chapter. All we need for the present argument is an important theorem about this transform called Parseval's theorem (Oppenheim and Schafer, 1989):

$$\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \cdot \sum_{k=0}^{N-1} |(\tilde{X}[k])|^2 \tag{6.8}$$

If we divide equation (6.8) by $1/N$, the left side is than formally equivalent to the variance of $x[n]$. The right side is the area under what is called the periodogram estimate of the power spectrum $P(f_k)$ with $f_k$ defined for discrete frequencies between zero and the Nyquist frequency (see Press et al., 1992, p. 551). In other words

$$s^2 = \sum_{f_k=0}^{F_{nyq}} P(f_k) \tag{6.9}$$

Since we can assume that the quantization noise is white, its power will be equally distributed between zero and the Nyquist frequency. Using both (6.5) and (6.9), we learn that the total area covered by quantization noise in the power spectrum is proportional to $Q^2/12$. Hence, the expected power spectral amplitude $\bar{P}_{quant}(f)$ for the quantization noise at frequency $f$ is proportional to the inverse of the Nyquist frequency $f_{Nyq}$

$$\bar{P}_{quant}(f) \sim \frac{1}{f_{Nyq}} \tag{6.10}$$

Thus, if we sample the same signal at two different sampling rates, the quantization noise levels will be related as the inverse of the Nyquist frequencies, while the total noise areas will be equal. This situation is sketched in Fig. 6.11.

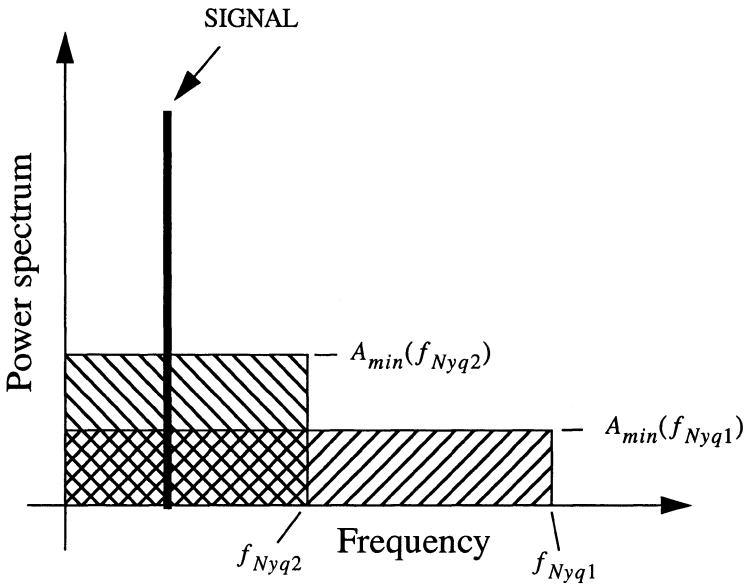



**Fig. 6.11** Signal and quantization noise level for different sampling frequencies.

To complete the argument we make use of the general definition of dynamic range given by equation (6.2). We assume that the amplitudes to be measured are power spectral amplitudes as shown in Fig. 6.11 and that $A_{max}$, the maximum measurable amplitude, corresponds to the spectral amplitude of what is sketched as SIGNAL in Fig. 6.11. The minimum measurable amplitude $A_{min}$ is assumed to be given by the noise level of the quantization noise. For the sake of this argument we can ignore other noise sources. Hence for the two different Nyquist frequencies $f_{Nyq1}$ and $f_{Nyq2}$, $A_{min}$ will correspond to $A_{min}(f_{Nyq1})$ and $A_{min}(f_{Nyq2})$ in Fig. 6.11, respectively. We can use (6.10) to write $A_{min}(f_{Nyq2})/A_{min}(f_{Nyq1}) \approx f_{Nyq1}/f_{Nyq2}$. Thus, if we initially sample the input signal at a much higher frequency than finally desired, the maximally achievable resolution for low amplitude signals will be larger than if we sample the input signal at the lower sampling frequency directly (Fig. 6.12 left and right panels, respectively). Crucial in this context is the fact that decimation does not involve any quantization and therefore does not change the noise level.

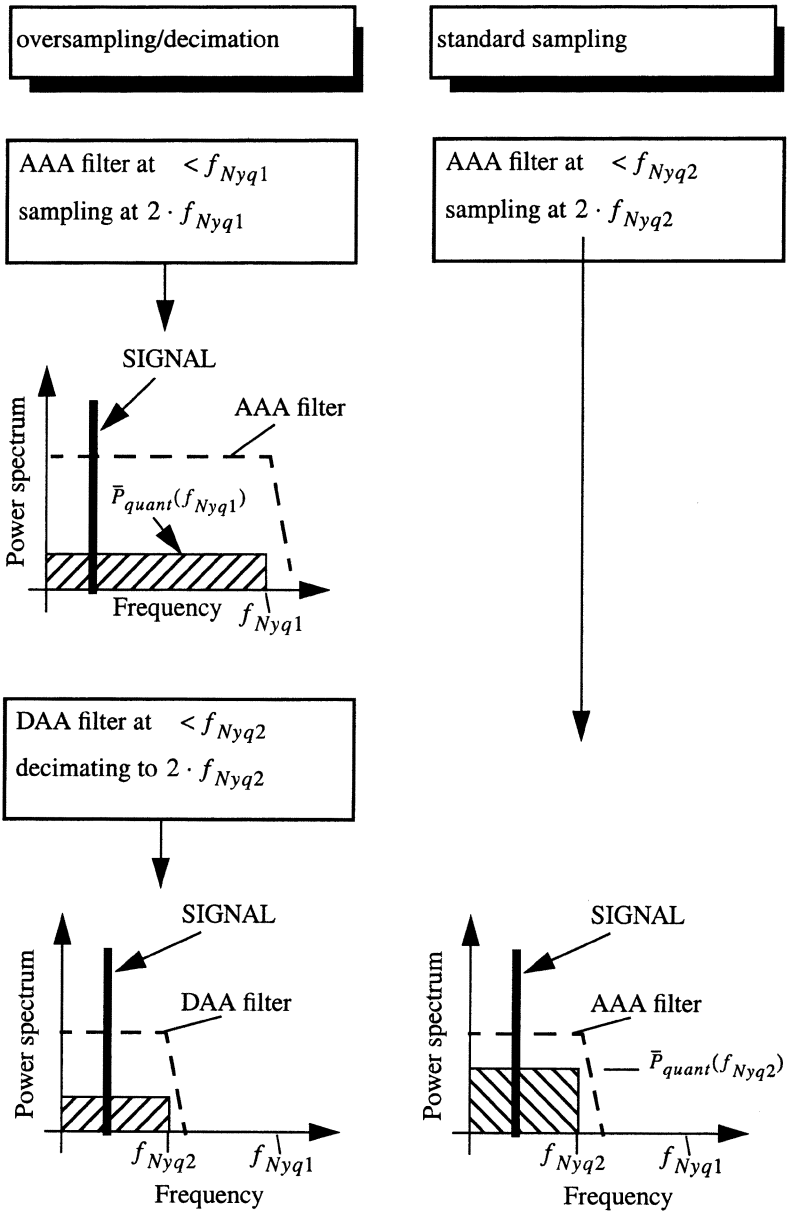



**Fig. 6.12** Comparison of oversampling/decimation and standard sampling. AAA filter is the analog anti-alias filter and DAA filter is the digital anti-alias filter.

The increase in resolution obtained for plain oversampling works on the assumption that the quantization noise is uniformly distributed between zero and the Nyquist frequency. A different set of oversampling methods employed in seismic recording systems is based on the quantization of amplitude differences between subsequent samples (differential sampling). For sampling frequencies much higher than the signal frequencies of interest, these changes are very small, especially if the trace is integrated before sampling. Therefore, ADCs with smaller numbers of bits (down to single-bit ADCs) suffice. This is exploited in sigma-delta modulation, which has become the most commonly used analog to digital conversion technique employed in modern digital recording systems (Steim, 1986). It involves the quantization of amplitude differences between subsequent samples of the integrated trace using a single bit ADC. Following the comprehensive description in Park (1999), we will discuss below the basic principles of this technique.

Sigma-delta modulation can be seen as an extension of the simpler delta modulation in which differential sampling is performed using a single-bit-quantizer as sketched in Fig. 6.13.

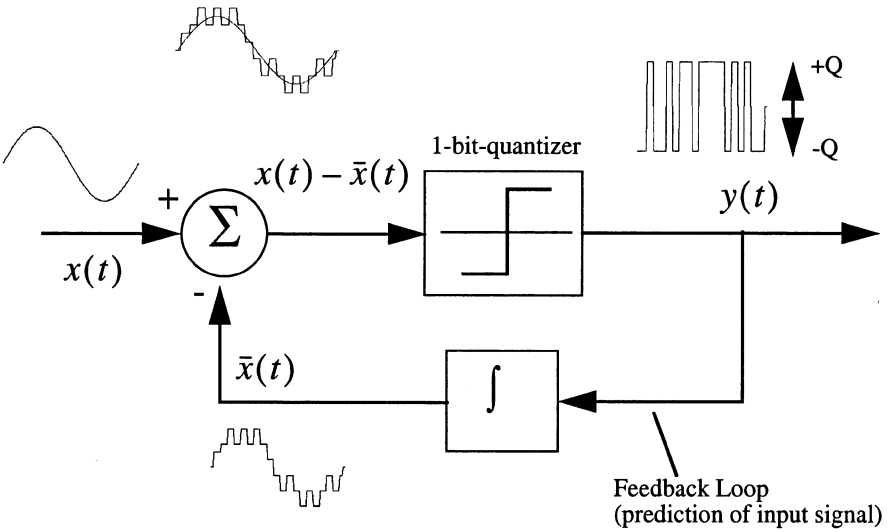



**Fig. 6.13** Principle of delta modulation.

The signal to be quantized is the deviation of the input signal $x(t)$ from a predicted input signal $\bar{x}(t)$. If this deviation $x(t) - \bar{x}(t)$ - which is called the prediction error - is positive, the single-bit-quantizer yields a positive output value $y(t) = +Q$. For negative output values on the other hand, a negative value $y(t) = -Q$ is obtained. In order to predict the input signal, the modulator output signal $y(t)$ is integrated in the feedback loop (Fig. 6.13).

Depending on the resolution of the 1-bit-quantizer, integrating the modulator output signal in Fig. 6.13 yields a fairly "rugged" signal. During demodulation this signal is smoothed using a lowpass filter as sketched in Fig. 6.14.

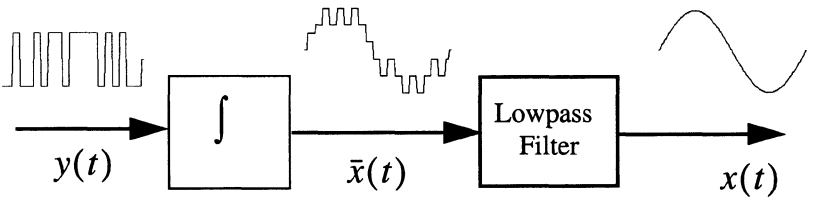



**Fig. 6.14** Principle of delta demodulation.

*Problem 6.3* What happens during delta modulation for a rapidly rising input signal? What happens if the input signal is constant? For this demonstration, generate a centered step function of amplitude 1. This is achieved within DST by integrating a spike of amplitude 100 generated using an internal sampling frequency of 100 Hz and a window length of 512 points. The centered step function contains both a rapidly rising and two constant parts. Apply the delta modulation simulation within DST using a LSB value of 0.01 and discuss the differences between actual input signal and predicted input signal.

As can be seen from Fig. 6.13 and Fig. 6.14, both delta modulation and demodulation require an integrator. The resulting demodulated output signal is the same if the integration is performed before the final lowpass filtering or before delta modulation (why?). This is demonstrated in Problem 6.4.

*Problem 6.4* Perform the two types of delta modulation discussed above on a sinusoidal test signal. For this demonstration, set the internal sampling frequency in DST to 256Hz and the window length to 512 points (*Modify* option of the *Setup* menu). Next, generate a sinusoidal input signal using the *Test Signals -> Sine/Cosine* option. For both the signal amplitude and the signal frequency enter a value of 1.0. First, simulate plain delta modulation using a LSB value of 0.05. Next, compare the resulting demodulated signal with the one you obtain if the integration is performed before delta modulation. In the latter case, adjust the LSB value for the quantization to the maximum peak to peak amplitude of the signal to be fed to the delta modulator. For the lowpass filter use a Butterworth filter (forward/backwards) using a single section and a corner frequency of 4 Hz.

The fact that integration can be performed either before modulation or before the final lowpass filtering without changing the resulting demodulated signal offers an intuitive way to see how plain delta modulation can be extended to sigma-delta modulation. This is schematically shown in Fig. 6.15.

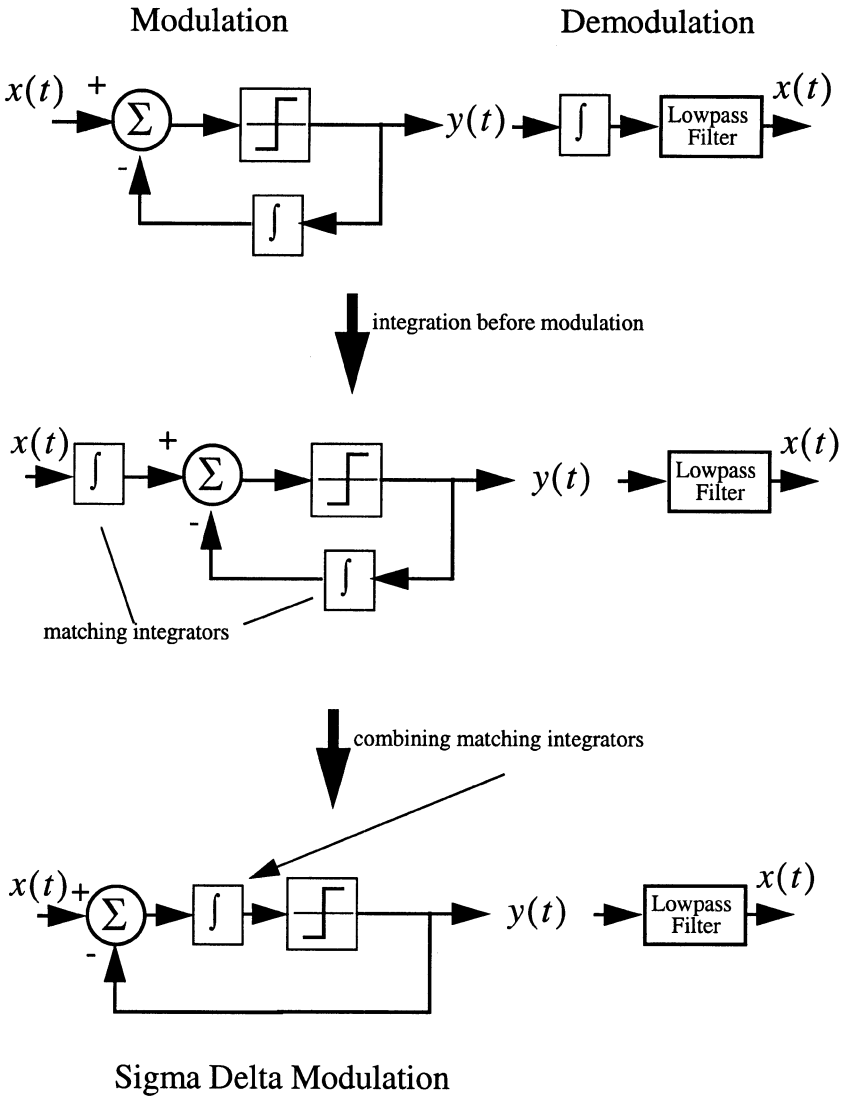



**Fig. 6.15** Derivation of sigma-delta modulation from delta modulation.

The uppermost panel in Fig. 6.15 shows again the principle of plain delta modulation and demodulation. If the integration before lowpass filtering is done before modulation, the block diagram shown in the middle panel of Fig. 6.15 is obtained. In this setup, the integration in the feedback loop matches the integration before modulation. Therefore, both integrators can be combined into a single one as shown in the bottom panel of Fig. 6.15. This way we arrive at the principle of sigma-delta modulation. The greek symbol *sigma*

($\Sigma$) stands for the integrator while the greek symbol *delta* ($\Delta$) stands for the delta modulator. In contrast to plain oversampling, for sigma-delta modulation the level of quantization noise is no longer independent of frequency. This can be seen from a simplified analysis of the s-plane transfer function (Park, 1999).

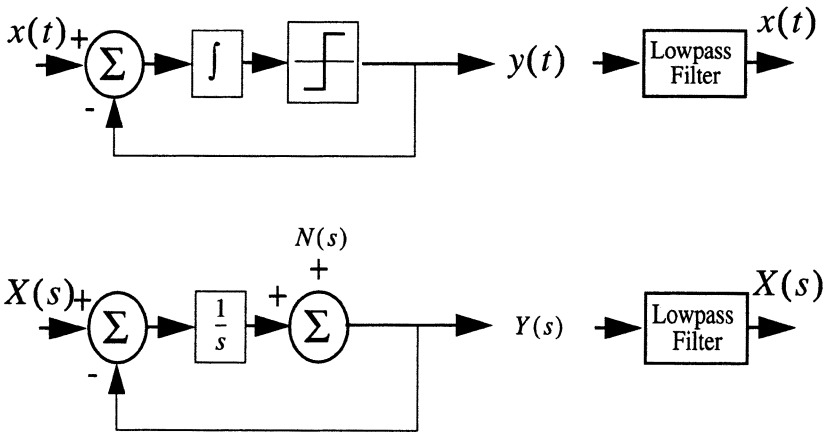



**Fig. 6.16** Simplified s-plane model of the sigma-delta modulator (after Park, 1999).

As can be seen from the block diagram in Fig. 6.16 (upper panel), the sigma-delta modulator quantizes the integrated difference between the input signal and the modulator output signal. Since integration corresponds to multiplication with $1/s$ in the s-plane, the Laplace transform of the sigma-delta modulator output is given by:

$$Y(s) = (X(s) - Y(s)) \cdot \frac{1}{s} + N(s) \tag{6.11}$$

$X(s)$ and $Y(s)$ are the Laplace transform of the modulator input and the output signal $x(t)$ and $y(t)$, respectively. The contribution of the quantizer is modelled by additive quantization noise $N(s)$. If $N(s) = 0$, the modulator output becomes

$$Y(s) = (X(s) - Y(s)) \cdot \frac{1}{s} \tag{6.12}$$

Hence, the signal transfer function becomes

$$\frac{Y(s)}{X(s)} = \frac{1}{s} \Big/ \left(1 + \frac{1}{s}\right) = \frac{1}{s+1} \tag{6.13}$$

On the other hand, the modulator output for $X(s) = 0$ is

$$Y(s) = -Y(s) \cdot \frac{1}{s} + N(s) \tag{6.14}$$

and the noise transfer function becomes

$$\frac{Y(s)}{X(s)} = 1 \Big/ \left(1 + \frac{1}{s}\right) = \frac{s}{s+1} \tag{6.15}$$

Since (6.13) and (6.15) correspond to a lowpass and a highpass filter, respectively, the sigma-delta modulator lowpass filters the signal and highpass filters the noise (noise shaping) as is schematically shown in Fig. 6.17

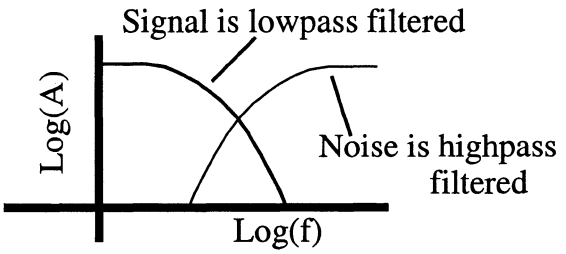



**Fig. 6.17** Noise shaping properties of a simplified model of sigma-delta modulation (after Park, 1999).

If sigma-delta modulation is combined with an analog-to-digital converter, we obtain a sigma-delta ADC. In this case, the sampled modulator output is converted back to an analog signal within the feedback loop (DAC in Fig. 6.18). If the input signal is strongly oversampled, the digital output signal can be averaged over several samples with a digital anti-alias filter and subsequently decimated to the desired sampling frequency. Although the individual samples carry a large quantization error, the averaging properties of the digital anti-alias filter (decimation filter in Fig. 6.18) yield a very precise result. For practical reasons, the decimation process is often realized as a sequence of individual stages each of which consists of a digital anti-alias filter combined with a decimation stage (see Fig. 6.18).

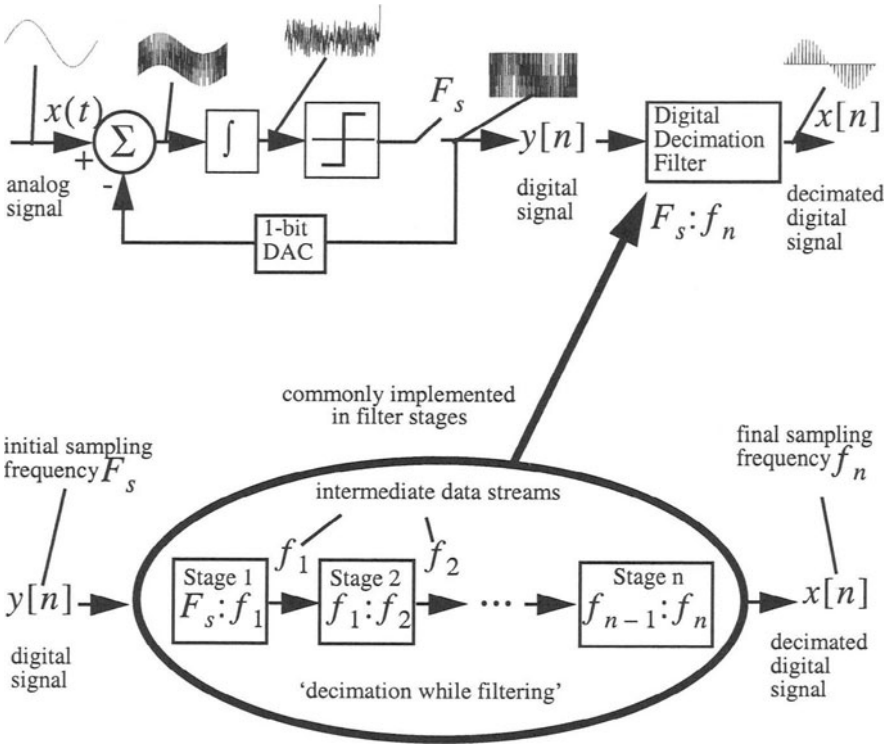



**Fig. 6.18** Principle of sigma-delta A/D conversion. $F_s$ indicates the initial sampling frequency, DAC denotes the digital-to-analog converter within the feedback loop, and $f_n$ denotes the final sampling frequency of the decimated digital signal.

Since for low frequencies the quantization noise level in sigma-delta modulation is generally smaller than for the plain oversampling technique, sigma-delta modulation is now commonly used as analog to digital conversion technique employed in modern digital recording systems (Steim, 1986). For more discussion of details of this technique the interested reader is referred to Park (1999) and the textbook of Proakis and Manolakis (1992).

*Problem 6.5* Use the DST to simulate sigma-delta analog-to-digital conversion on a 2 second long sinusoidal test signal of peak amplitude 1 and a signal frequency of 1 Hz. For the internal sampling frequency use 5120 Hz. Decimate the output signal of the sigma-delta modulator in 2 decimation stages. Use a decimation ratio of 16 for the first step followed by a decimation ratio of 4. Try to find the optimum LSB value.

Independent of the special flavour of the oversampling technique used, digital seismograms obtained from modern seismic acquisition systems can be seen as the output of a sequence of filter stages which involve both *analog* and *digital* components (Fig. 6.19). In addition to the advantages in terms of dynamic range and resolution, these systems allow instrument designers to use fairly simple analog anti-alias filters and do most of the work in the digital domain where powerful digital anti-alias filters are easy to implement.

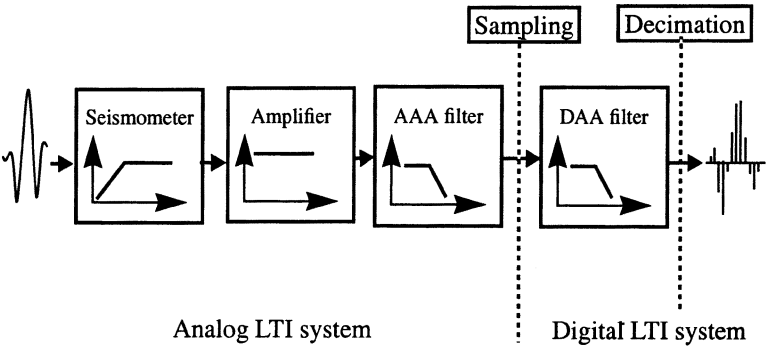



**Fig. 6.19** Building blocks of modern seismic acquisition systems. AAA filter denotes the analog anti-alias filter, whereas DAA filter is the digital anti-alias filter to be applied during decimation and is often implemented as a sequence of filter stages (e. g. Fig. 6.18).

CHAPTER 7

# From infinitely continuous to finite discrete

In previous chapters we developed concepts which describe the properties of continuous-time LTI systems, such as is shown in the analog part of Fig. 6.19. On the other hand, when we used DST to demonstrate these properties we were, of course, working with sequences of numbers. This was acceptable in the context of the examples we used. In a general context, however, we must be well aware of several important differences between discrete and continuous systems. Our intuitive grasp of the more important system properties must by now be firm enough that we can extend our view using a more formal approach. In this fashion we will not only acquire additional tools for data processing, but also we will gain some insight into the links between systems defined for *infinite continuous-time signals* and systems defined for *finite discrete-time signals*. The focus of this chapter will be on consequences of this transition for the concepts which we developed so far. We will see for example that the transition from continuous-time to discrete-time systems corresponds to a transition from aperiodic to periodic Fourier spectra. This property for example will allow us to take a new look at the aliasing problem. Furthermore, we will meet the z-transform for discrete sequences, the discrete counterpart of the Laplace transform. We will see that the concepts introduced for continuous-time systems, such as the transfer function or the concept of poles and zeros, remains valid for discrete-time systems as well. Even more important, we will see that most of the relevant system properties can intuitively be understood from their analog counterparts. In the following the sequence notation of Oppenheim and Schafer (1989) will be used.

## 7.1 Fourier series for periodic continuous-time signals (CTFS)

Let us start out with some of the basic properties of Fourier analysis named after the french mathematician Jean Baptiste Joseph Fourier (1768-1830). Fourier's work on trigonometric series expansions of the temperature distribution through bodies led him to suggest that trigonometric series representations could be used for any periodic function. The conditions under which this is valid were defined by Peter Gustav Lejeune Dirichlet (1805-1859) more than 20 years after Fourier's original proposal. These conditions are now known as Dirichlet conditions and are fulfilled for any periodic function $x_p(t)$ if

- it has only a finite number of discontinuities
- it contains only a finite number of extreme values and
- it is absolutely integrable, that is $\int_{T_0} |x_p(t)| < \infty$

for any period $T_0$.

Essentially all periodic signals $x_p(t)$ of practical interest meet these requirements and can be described by a Fourier series

$$x_p(t) = \sum_{k=-\infty}^{\infty} c_k \cdot e^{j \cdot k\frac{2\pi}{T_0} \cdot t} = \sum_{k=-\infty}^{\infty} c_k \cdot e^{j \cdot k2\pi f_0 \cdot t}. \tag{7.1}$$

$T_0$ is the period of $x_p(t)$ and is also called the fundamental period. Its reciprocal $f_0 = 1/T_0$ is called the fundamental frequency. The complex Fourier coefficients $c_k$ are given by

$$c_k = \frac{1}{T_0}\int_{T_0} x_p(t) \cdot e^{-j \cdot k2\pi f_0 \cdot t} dt \tag{7.2}$$

An alternative formulation of the Fourier series for real signals $x_p(t)$ in terms of sine and cosine series is given by

$$x_p(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(k2\pi f_0 \cdot t) + \sum_{k=1}^{\infty} b_k \sin(k2\pi f_0 \cdot t) \tag{7.3}$$

with

$$\begin{aligned} a_0 &= c_0 \\ a_k &= 2 \cdot Re\{c_k\} \\ b_k &= -2 \cdot Im\{c_k\} \end{aligned} \tag{7.4}$$

The derivation of (7.3) from (7.1) can be found in Strum and Kirk (1988). The Fourier series representation expresses a periodic continuous-time signal as a linear combination of harmonic signals. In this context, the Fourier series can be thought of as mapping a periodic 'time' signal $x_p(t)$ from the 'time domain' into a 'frequency domain' defined for discrete frequencies which are integer multiples of the fundamental frequency $f_0$. Of course, the meaning of 'time-' and 'frequency domains' are not restricted to physical time and frequency.

### 7.2 Fourier transform for aperiodic continuous-time signals (CTFT)

For aperiodic continuous-time signals $x(t)$ a frequency domain representation similar to equation (7.2) is given by the Fourier transform $\mathcal{F}\{x(t)\}$ which is defined by the following equation:

$$\mathcal{F}\{x(t)\} = X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}\, dt. \tag{7.5}$$

Heuristically, this representation can be derived from the integral in (7.2) by considering a periodic signal $x_p(t)$ with period $T_0$ in the limit as $T_0 \to \infty$. In this case, the limits of the integral approach $\infty$ and the discrete frequencies $k\frac{2\pi}{T_0} = k2\pi f_0 \to 2\pi f$. A sufficient set of conditions for the existence of the Fourier transform is again given by the Dirichlet conditions for $T_0 \to \infty$.

It is common practice to write the frequency in terms of angular frequency $\omega = 2\pi f$, and since (7.5) describes a complex quantity to write:

$$X(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\, \mathrm{dt} \tag{7.6}$$

For the inverse transformation we have:

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\omega)e^{j\omega t}\, d\omega \tag{7.7}$$

One reason for using $j\omega$ is that this allows us right away to see the Fourier transform as a special case of the Laplace transform evaluated on the imaginary axis ($j\omega$) of the complex s-plane (cf. chapter 2.5).

The Fourier transform contains certain symmetry properties which may be used to increase the computational efficiency but which are also helpful in understanding some properties of linear filters, e. g. the time domain symmetry of zero phase filters discussed in chapter 8. If a signal has certain symmetry properties in one domain such as being even or odd, etc. its transform will have corresponding symmetry properties. Some of these correspondences are listed in Tab. 7.1.

**Table 7.1** Symmetry properties of the Fourier transform.

| time domain | frequency domain |
|---|---|
| x(t) real | X(-jω) = [X(jω)]* |
| x(t) imaginary | X(-jω) = -[X(jω)]* |
| x(t) = x(-t) (x(t) even) | X(-jω) = X(jω) (X(jω) even) |
| x(t) = - x(-t) (x(t) odd) | X(-jω) = -X(jω) (X(jω) odd) |
| x(t) real and even | X(jω) real and even |
| x(t) real and odd | X(jω) imaginary and odd |
| x(t) imaginary and even | X(jω) imaginary and even |
| x(t) imaginary and odd | X(jω) real and odd |
| | * denotes complex conjugate |

Of additional importance in the present context are the Fourier transform properties for certain mathematical operations such as convolution, time shifting, derivation, and integration which were introduced in chapter 2.3.

### 7.3 Fourier transform of discrete-time signals

In the case of digital seismic signals, we are dealing with discrete-time signals which are obtained by taking samples from continuous signals at discrete times $nT$ with $T$ being the sampling interval in sec. We can view this process as being caused by the multiplication of the continuous signal with a periodic impulse train

$$\delta_T(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT) \tag{7.8}$$

with $\delta(t - nT)$ being time shifted delta functions. If the continuous time signal would be $x_c(t)$, the sampled signal can be written

$$x_s(t) = x_c(t)\delta_T(t) \tag{7.9}$$

In order to more formally describe the consequences of sampling in the frequency domain, we use the following two Fourier transform properties for which the proofs can be found in Kraniauskas (1992) ( $\Leftrightarrow$ indicates a transform pair, $x(t) \Leftrightarrow X(j\omega)$ ):

$$\bullet\ \textit{Multiplication} — \ x(t) \cdot h(t) \Leftrightarrow \frac{1}{2\pi} X(j\omega) * H(j\omega) \tag{7.10}$$

$$\bullet\ \textit{Impulse train transform} — \ \sum_{n=-\infty}^{\infty} \delta(t - nT) \Leftrightarrow \frac{2\pi}{T} \sum_{k=-\infty}^{\infty} \delta\left(\omega - k\frac{2\pi}{T}\right) \tag{7.11}$$

Using (7.10) we can write the Fourier transform of $x_s(t)$ as the convolution ($*$) of the Fourier transforms of $\delta_T(t)$ and $x_c(t)$ which are denoted $\Delta_T(j\omega)$ and $X_c(j\omega)$, respectively. From (7.11) it follows that

$$\Delta_T(j\omega) = \frac{2\pi}{T} \sum_{k=-\infty}^{\infty} \delta(\omega - k\omega_s) \tag{7.12}$$

with $\omega_s = 2\pi/T$ being the sampling frequency in rad/sec. Convolving $\Delta_T(j\omega)$ and $X_c(j\omega)$ yields

$$X_s(j\omega) = \frac{1}{T} \sum_{k=-\infty}^{\infty} X_c(j\omega - kj\omega_s) \tag{7.13}$$

Hence the Fourier transform of a sampled signal is periodic with sampling frequency $\omega_s$. If the sampled signal does not contain energy at and above $\omega_s/2$ that is if the sampling is done in accordance with the sampling theorem in chapter 5, the Fourier transform of the continuous signal $X_c(j\omega)$ can be obtained from a single period of the Fourier transform of the sampled signal by

$$X_c(j\omega) = T \cdot X_s(j\omega) \tag{7.14}$$

This situation is illustrated in Fig. 7.1 a and b. If the sampling theorem is violated, however, that is if the sampling frequency is too small in comparison to the frequency content of the signal, it is no longer possible to obtain $X_c(j\omega)$ from $X_s(j\omega)$. This results in the alias effect shown in Fig. 7.1c.

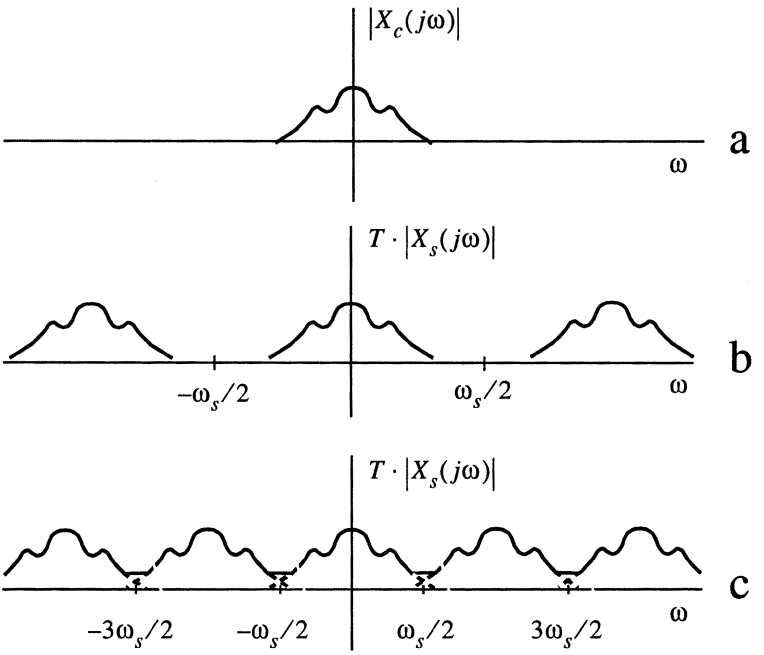



**Fig. 7.1** The effect of sampling in the frequency domain. Fig. 7.1a shows a generic Fourier amplitude spectrum for a band-limited continuous-time signal. Fig. 7.1b shows the corresponding spectrum for the sampled signal in which sampling has been performed in accordance with the sampling theorem. Fig. 7.1c shows the corresponding spectrum for the sampled signal if the sampling theorem is violated (alias effect). Amplitudes are scaled according to (7.14).

Going back to the definition of the Fourier transform for aperiodic continuous-time signals, we can derive the transform for sampled signals if we evaluate

$$\begin{aligned}\mathcal{F}\{x(t)\delta_T(t)\} &= \int_{-\infty}^{\infty} x(t)\delta_T(t)e^{-j2\pi ft}\,dt \\ &= \int_{-\infty}^{\infty} x(t)\left(\sum_{n=-\infty}^{\infty}\delta(t-nT)\right)e^{-j2\pi ft}\,dt \\ &= \sum_{n=-\infty}^{\infty} x(nT)e^{-j2\pi fnT}\end{aligned} \tag{7.15}$$

Making the transition from the continuous-time signal $x(t)$ sampled at $nT$ to the sequence $x[nT]$ - by formally replacing $x(nT)$ by $x[nT]$ - we obtain the Fourier transform for discrete-time sequences

$$\mathcal{F}\{x[nT]\} = X_d(j\omega) = \sum_{n=-\infty}^{\infty} x[nT]e^{-j\omega nT} \tag{7.16}$$

Here, $x[nT]$ denotes a discrete-time sequence for which the time interval between individual samples is $T$ and $\omega = 2\pi f$ is the angular frequency. We have just made the important transition from continuous to discrete systems. Mathematically this corresponds to the transition from functions defined for real domains (denoted by ()) to functions whose domains are integer (denoted by []).

Above it was demonstrated that the sampling process introduces a periodicity with $\omega_s = 2\pi/T$ into the corresponding spectrum (cf. equation (7.13)). With the definition of the Fourier transform for discrete-time signals in (7.16) this effect can also be seen by replacing $\omega$ by $\omega + 2\pi/T$ which results in $X_d\left(j\left(\omega + \frac{2\pi}{T}\right)\right) = X_d(j\omega)$ . More general $X_d\left(j\left(\omega + \frac{2\pi}{T}r\right)\right) = X_d(j\omega)$ for any integer $r$. Therefore, in the discrete-time signal $x[nT]$ frequencies of $\omega$ and $\omega + \frac{2\pi}{T}r$ can not be distinguished and for the synthesis we need to consider only a frequency range of length $2\pi/T$.

This leads to the inverse transform (Kraniauskas, 1992):

$$\mathcal{F}^{-1}\{X_d(j\omega)\} = x[nT] = \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T} X_d(j\omega)e^{j\omega nT}d\omega \tag{7.17}$$

Equations (7.16) and (7.17) form a Fourier transform pair. Since the frequency response function of a linear system equals the Fourier transform of the impulse response, *the frequency response function of a discrete-time system is always a periodic function with a period of* $\omega_s = 2\pi/T$.

It should be noted that in some textbooks the Fourier transform for discrete-time sequences and its inverse transform are given in a different notation which deserves some comments. For example Oppenheim and Schafer (1989) use the form

$$\mathcal{F}\{x[n]\} = X(e^{j\tilde{\omega}}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\tilde{\omega}n} \tag{7.18}$$

for the Fourier transform of discrete-time sequences and

$$\mathcal{F}^{-1}\{X(e^{j\tilde{\omega}})\} = x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\tilde{\omega}})e^{j\tilde{\omega}n}d\omega \tag{7.19}$$

for the inverse transform. Here, $x[n]$ replaces $x[nT]$ emphasizing the fact that $x[n]$ is merely a sequence of numbers which may or may not have been obtained by a physical sampling process. Furthermore, $\omega T$ is replaced by the digital frequency[1]

$$\tilde{\omega} = \frac{\omega}{F_{samp}} = \omega T \tag{7.20}$$

with the sampling frequency $F_{samp}$. In this notation the spectrum $X(e^{j\tilde{\omega}})$ becomes periodic with a period of $2\pi$. The purpose of writing $X(e^{j\tilde{\omega}})$ is to indicate the close relationship between the Fourier transform for discrete sequences and the z-transform which will be discussed below. Digital frequency is often used in textbooks dealing exclusively with discrete-time processing while in books which deal with both analog and digital systems, the notation of (7.16) and (7.17) is often preferred. Here, $\omega$ will be

1. Oppenheim and Schafer (1989) use $\omega$ for digital frequency which has been replaced by $\tilde{\omega}$ in this text to avoid confusion.

used for analog frequency and $\tilde{\omega}$ for digital frequency.

It is also interesting to note that in contrast to the continuous-time case, there is an asymmetry between equations (7.16) and (7.17). This corresponds to a modification of the equivalence between convolution and multiplication. While convolution of sequences corresponds to multiplication of periodic Fourier transforms, multiplication of sequences corresponds to periodic convolution (with the convolution only carried out over one period) of the corresponding Fourier transforms.

### 7.4 Fourier series for periodic discrete-time signals

In case a sampled discrete-time signal is periodic with period $T_0 = NT$, that is $x_p[nT] = x_p[nT + NT]$ for all $n$, the Fourier series representation becomes

$$c_k = \frac{1}{T_0}\int_{T_0}\left(\sum_{n=-\infty}^{\infty}\delta(t-nT)\right)x_p(t)\cdot e^{-j\cdot k2\pi f_0\cdot t}dt. \tag{7.21}$$

Due to the sifting property of the delta function, the integral over $T_0$ contributes only at $N$ discrete times $nT$ for $n = 0, ...,$N-1 and (7.21) becomes

$$\begin{aligned} c_k &= \frac{1}{NT}\sum_{n=0}^{N-1} x_p(nT)\cdot e^{-j\cdot k\frac{2\pi}{NT}\cdot nT} \\ &= \frac{1}{NT}\sum_{n=0}^{N-1} x_p[nT]\cdot e^{-j\cdot k\frac{2\pi}{N}\cdot n} \end{aligned} \tag{7.22}$$

We have already noted for continuous-time signals that only discrete frequencies (here $\omega_k = k\frac{2\pi}{NT}$) contribute to the Fourier series representation of periodic signals. In addition however, we have seen that the spectra of discrete-time sequences are periodic with $\omega_s = 2\pi/T$. In the present context this means that the $c_k$ repeat themselves with period $N$ ($c_k = c_{k+N}$ for all $k$). Since for the synthesis of discrete signals we only need to consider a frequency range of length $2\pi/T$ (cf. (7.17)), only the discrete frequencies $\omega_k$ for $k = 0, ..., N-1$ contribute to the Fourier series synthesis equation for discrete-time periodic signals which becomes:

$$x_p[nT] = \sum_{k=0}^{N-1} c_k \cdot e^{j \cdot k \frac{2\pi}{NT} \cdot nT} = \sum_{k=0}^{N-1} c_k \cdot e^{j \cdot k \frac{2\pi}{N} \cdot n} . \tag{7.23}$$

## 7.5 The Discrete Fourier Transform (DFT)

So far we are still considering sequences of infinite duration. In reality, of course, we are always dealing with finite sequences for which a different Fourier representation - the Discrete Fourier Transform (DFT) - is commonly used. In contrast to the Fourier transform for infinite sequences, it is not a continuous function but itself a finite length sequence. For a finite discrete-time sequence $x[nT]$ of length $N$, the DFT can be defined as:

$$\tilde{X}[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[nT] e^{-j2\pi kn/N} \tag{7.24}$$

The set of sampled values can be recovered from the DFT by means of the inverse DFT, given by:

$$\tilde{x}[nT] = \sum_{k=0}^{N-1} \tilde{X}[k] e^{j2\pi kn/N} \tag{7.25}$$

You can see from equation (7.25) that the inverse DFT yields the periodic sequence $\tilde{x}[nT]$ from which $x[nT]$ can be obtained by considering only a single period. If the length of the sequence $x[nT]$ equals an integer power of 2, the computation of the DFT and inverse DFT is normally done using the time efficient FFT (Fast Fourier Transform) algorithm (Oppenheim and Schafer, 1989). If not, the sequences are normally padded with zeros up to the next integer power of 2.

Except for the scaling factor of $1/T$ the DFT it is equivalent to the Fourier series representation (equation (7.22)) of the infinite periodic sequence $\tilde{x}[nT]$ which is made up by periodic extension of the given finite sequence $x[nT]$

$$\tilde{X}[k] = c_k \cdot T \tag{7.26}$$

Here $\tilde{x}[nT]$ is the (infinite) periodic sequence constructed from $x[nT]$ by periodic continuation. As we will see below when the convolution of finite sequences is discussed, in

the context of the DFT it saves a lot of trouble to always think in terms of $\tilde{x}[nT]$, the periodic extension of $x[nT]$ and consider $x[nT]$ as just one period of $\tilde{x}[nT]$. The DFT is only defined for discrete radian frequencies $\omega_k$, which are related to the total number of points $N$, and the sampling interval $T$ by:

$$\omega_k = k \cdot \frac{2\pi}{TN} \quad \text{for k} = 0, 1, \ldots - \ldots, N-1 \tag{7.27}$$

Or in terms of frequencies $f_k$

$$f_k = k \cdot \frac{1}{TN} = k \cdot \frac{f_{dig}}{N} \quad \text{for k} = (0, 1, \ldots - \ldots, N-1) \tag{7.28}$$

with $f_{dig}$ being the sampling frequency.

When comparing the representation of arbitrary continuous-time signals of infinite duration and discrete-time signals of finite duration in the spectral domain, we notice that the corresponding 'Fourier spectra' (equations (7.6) and (7.24)) have changed from being aperiodic in a continuous-frequency domain to being periodic for discrete frequencies. The transition from continuous-time to discrete-time by sampling corresponds to the transition from aperiodic to periodic spectra, while the transition from aperiodic signals to periodic signals corresponds to the transition from continuous spectra to discrete spectra. This is schematically displayed in Fig. 7.2.

**Continuous-time Fourier transform (CTFT)**
*signal:* infinite
*time:* continuous
*spectrum:* aperiodic
*frequency:* continuous

**Continuous-time Fourier series (CTFS)**
*signal:* infinite, periodic
*time:* continuous
*spectrum:* aperiodic
*frequency:* discrete

**Discrete-time Fourier transform (DTFT)**
*signal:* infinite
*time:* discrete
*spectrum:* periodic
*frequency:* continuous

**Discrete Fourier transform (DFT)**
*signal:* finite
*time:* discrete
*spectrum:* periodic
*frequency:* discrete

**Fig. 7.2** Properties of continuous-time and discrete-time signals and their corresponding Fourier spectra.

In the context of defining the DFT and its inverse transform in (7.24) and (7.25) the particular notation $x[nT]$ was chosen based on the view of $x[nT]$ as a finite sequence of samples of a discrete-time process. In addition it conveniently helped to demonstrate the systematic changes of the spectral properties that accompany the transition from continuous-time to discrete-time systems. It should be noted, however, that in the context of the numerical implementation, the DFT is most commonly defined for mere sequences of numbers and the notation $x[n]$ is used instead of $x[nT]$ (e.g. Strum and Kirk, 1988; Oppenheim and Schafer, 1989). Furthermore, the scaling factor $1/N$ is most commonly kept with the inverse transform. Therefore, a definition commonly found in software packages (e.g. IEEE Digital Signal Processing Committee, 1979) is:

$$\tilde{\bar{X}}[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N} \tag{7.29}$$

for the DFT and

$$\tilde{\tilde{x}}[n] = \frac{1}{N}\sum_{k=0}^{N-1} \tilde{X}[k]e^{j2\pi kn/N} \tag{7.30}$$

for the inverse DFT. This directly leads to a common practical problem.

*Problem 7.1* Given a finite sequence of numbers $x[n]$ ($n = 0$ to $N-1$), which was obtained by sampling of a infinite continuous-time signal $x(t)$ (sampling interval $T$ sec), what is the relationship between the DFT given in the notation of (7.29) and the Fourier transform given by (7.6) provided a) that the sampling was done in accordance with the sampling theorem, and b) that the amplitude of the signal $x(t)$ is zero outside of the finite time window ($0 \le t < NT$)? Hint: Use the rectangular rule to approximate the Fourier transform integral and compare the resulting finite sum for radian frequencies $\omega_k = k \cdot \frac{2\pi}{TN}$ to the expression for the DFT.

## 7.6 The z-transform and the discrete transfer function

While the analog component of the recording system shown in Fig. 6.19 can be described in terms of its analog transfer function $T(s)$, for the digital component there is a discrete counterpart, the discrete transfer function $T(z)$. While $T(s)$ was defined in terms of the Laplace transform, $T(z)$ is defined in terms of its discrete counterpart, the z-transform.

The bilateral z-transform of a discrete sequence $x[n]$ is defined as

$$\mathcal{Z}\{x[n]\} = \sum_{n=-\infty}^{\infty} x[n]z^{-n} = X(z) \tag{7.31}$$

The z-transform transforms the sequence $x[n]$ into a function $X(z)$ with $z$ being a continuous complex variable. With this definition we can define the discrete transfer function as the z-transform of the output $\mathcal{Z}\{y[n]\}$ divided by the z-transform of the input $\mathcal{Z}\{x[n]\}$ of a discrete linear system.

$$T(z) = \frac{\mathcal{Z}\{y[n]\}}{\mathcal{Z}\{x[n]\}} \tag{7.32}$$

To see the relationship between the z-transform and the Laplace transform for discrete signals, we go back to the definition of the Laplace transform (2.25) and apply it to a sampled signal.

$$\begin{aligned}\mathcal{L}\{x(t)\delta_T(t)\} &= \int_{-\infty}^{\infty} x(t)\delta_T(t)e^{-st}\,dt \\ &= \int_{-\infty}^{\infty} x(t)\left(\sum_{n=-\infty}^{\infty}\delta(t-nT)\right)e^{-st}\,dt \\ &= \sum_{n=-\infty}^{\infty} x(nT)e^{-snT}\end{aligned} \tag{7.33}$$

Making the transition from the continuous-time signal $x(t)$ sampled at $nT$ to the sequence $x[nT]$ - by formally replacing $x(nT)$ by $x[nT]$ - we obtain the Laplace transform for discrete-time sequences

$$\mathcal{L}\{x[nT]\} = \sum_{n=-\infty}^{\infty} x[nT]e^{-snT} \tag{7.34}$$

Here, $x[nT]$ denotes a discrete-time sequence for which the time interval between individual samples is $T$. If we define $z = e^{sT}$ and $x[n] = x[nT]$ we obtain (7.31), the definition of the z-transform.

We can visualize how the complex s-plane is mapped onto the z-plane if we substitute $z = e^{sT} = e^{\sigma T}e^{j\omega T} = re^{j\omega T}$. This is sketched in Fig. 7.3. The origin ( $s = 0$ ) maps onto $z = 1$ while $s = j\omega_s/2 = j\pi/T$ maps onto $e^{j\pi} = \cos(\pi) = -1$. Since $z = e^{\sigma T}e^{j\omega T} = re^{j\omega T}$, the left half s-plane (the region with $s = \sigma + j\omega$ for $\sigma < 0$) maps to the interior of the unit circle in the z-plane ($r < 1$ because $e^{|\sigma|T} > 1$ and $r = e^{\sigma T} = e^{-|\sigma|T} = 1/e^{|\sigma|T}$). Likewise, the right half s-plane is mapped onto the outside of the unit circle. For points on the imaginary axis ($s = j\omega$), $r$ is 1 and z becomes the unit circle. Positive frequencies are mapped onto the upper half unit circle and negative frequencies onto the lower half. The complete linear frequency axis is wrapped around the unit circle with $\omega_s/2$ + multiples of $2\pi$ being mapped onto $z = -1$ (Fig. 7.3). For points on the unit circle ($z = e^{j\omega T}$), equation (7.31) reduces to the Fourier

transform for discrete-time sequences $x[n]$ (see equation (7.16)). In other words

> • The Fourier transform for discrete-time sequences equals the z-transform evaluated on the unit circle.

This analogy to the continuous-time case, where the Fourier transform equals the Laplace transform evaluated on the imaginary axis, reiterates the statement that the z-transform is the discrete counterpart of the Laplace transform. Most of the properties of discrete systems can intuitively be understood from the properties of analog systems if we replace the 'role' of the imaginary axis in the continuous-time case by the 'role' of the unit circle in the discrete-time case.

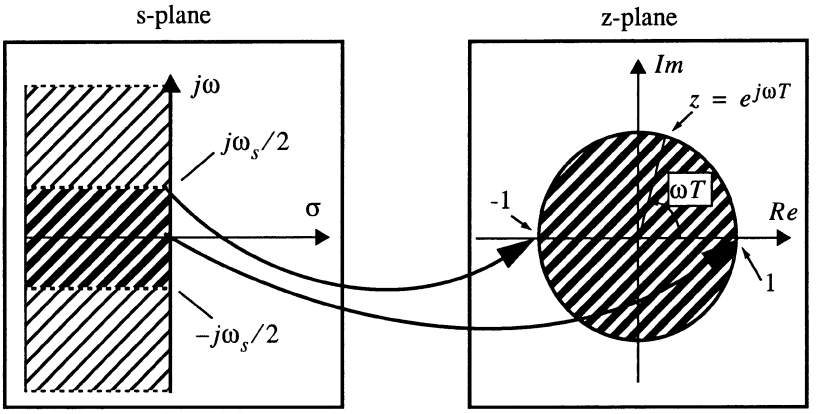



**Fig. 7.3 :** Mapping of the s-plane to the z-plane.

Like the Laplace transform, equation (7.31) does not necessarily converge for all values of $z$. Those regions where it does exist are called regions of convergence. For continuous-time systems we have seen that we can directly relate most of the essential system properties to the distribution of the system's poles and zeros in the s-plane. The same is true for the discrete-time system. Moreover, we can directly transfer the relationships between the pole-zero distribution in the s-plane to draw conclusions about the system properties to the z-plane if the singularities (poles and zeros) are mapped according to $z = e^{sT}$ (Fig. 7.3). Singularities of transfer functions which lie in the left half s-plane are mapped to singularities inside the unit circle in z while singularities in the right half-plane map to the outside of the unit circle (Fig. 7.4).

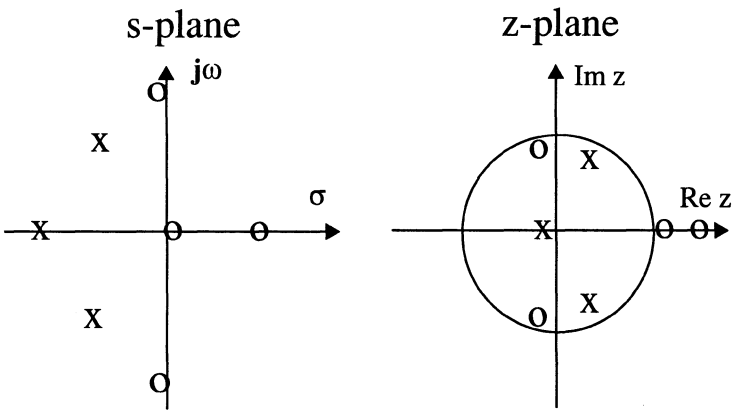



**Fig. 7.4** Pole-zero distribution in s-plane and z-plane. The poles/zeros in the s-plane close to $j\omega$ will be mapped close to the unit circle. Poles/zeros on the real axis will stay on the real axis.

Most of the transform properties for the z-transform can be directly deduced from their Laplace transform counterparts. The two most important z-transform properties in the present context are the shifting theorem

$$( x[n-n_0] \Leftrightarrow z^{-n_0}X(z) ). \tag{7.35}$$

and the convolution theorem

$$( x_1[n] * x_2[n] \Leftrightarrow X_1(z) \cdot X_2(z) ) \tag{7.36}$$

In this context the convolution of two sequences is defined as

$$x_1[n]*x_2[n] = \sum_{m=-\infty}^{\infty} x_1[m]x_2[n-m] \tag{7.37}$$

A property which becomes important in the context of chapter 8, is that replacing $z$ by $1/z$ in the z-transform ( $X(z)$ ) corresponds to inverting the input signal ( $x[n]$ ) in time

$$( x[-n] \Leftrightarrow X(1/z) ). \tag{7.38}$$

Hence from what we learned about the Laplace transform for continuous systems, we can intuitively justify the following properties for discrete systems:

- The region of convergence is a ring centered on the origin.
- The Fourier transform of $x[n]$ converges absolutely if and only if the unit circle is part of the region of convergence of the z-transform of $x[n]$ .
- The region of convergence must not contain any poles.
- If $x[n]$ is a sequence of finite duration, the region of convergence is the entire z-plane except possibly $z = 0$ or $z = \infty$.
- For a right-sided sequence the region of convergence extends outwards from the outermost pole to (and possibly including) $z = \infty$.
- For a left-sided sequence the region of convergence extends inwards from the innermost pole to (and possibly including) $z = 0$.
- For a two-sided sequence the region of convergence will consist of a ring bounded by poles on both the interior and the exterior.

These properties are further illustrated in Fig. 7.5.

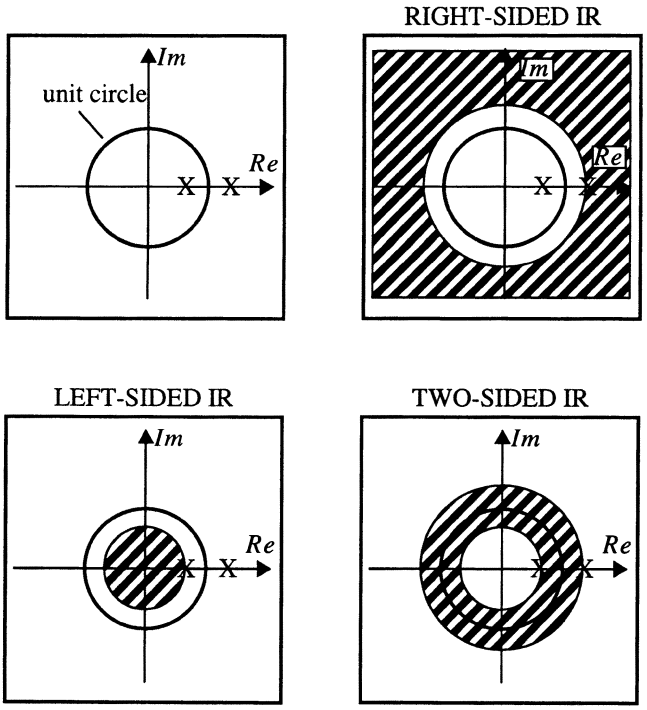



**Fig. 7.5** Region of convergence and type of resulting impulse response functions (IR) for a discrete system containing two poles on the real axis (cf. also Fig. 3.1).

Similar to the properties above, we can intuitively understand that for a discrete system to be **causal** and **stable**, all poles must lie inside the unit circle. If, in addition, the system contains no zeros outside the unit circle, it is called **minimum phase** system.

*Problem 7.2* Which type of impulse response function would result if we want to use the inverse Fourier transform for evaluation of a system containing two poles on the real axis as shown in Fig. 7.5?

*Problem 7.3* How do time shifts effect the region of convergence of the z-transform? Argue by using the shifting theorem for z-transforms ( $x[n-n_0] \Leftrightarrow z^{-n_0}X(z)$ ) for positive and negative $n_0$.

## 7.7 The inverse z-transform

When we need to evaluate the inverse z-transform, that is in order to obtain the impulse response function for a given discrete transfer function, we can proceed in several ways. Formally, we have to evaluate a contour integral

$$x[n] = \frac{1}{2\pi j} \oint_C X(z) z^{n-1} dz \tag{7.39}$$

with the closed contour C lying within a region of convergence and being evaluated counterclockwise. If the region of convergence includes the unit circle and if (7.39) is evaluated on it, the inverse z-transform reduces to the inverse Fourier transform (e.g. see Oppenheim and Schafer, 1989). This can be intuitively understood from the fact that the Fourier transform for discrete-time sequences equals the z-transform evaluated on the unit circle. Besides the direct evaluation of (7.39), the inverse z-transform can be obtained by power series expansion or partial fraction expansion into simpler terms for which the inverse z-transforms are tabulated and can be looked up in a table (Proakis and Manolakis, 1992). Partial fraction expansion is particularly useful if the z-transform is a rational function in $z$ which is always the case for the type of discrete linear systems we are dealing with. The reason is that these can always be described by linear difference equations with constant coefficients:

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{l=0}^{M} b_l x[n-l] \tag{7.40}$$

We take the z-transform of (7.40) using the shifting theorem ($x[n-n_0] \Leftrightarrow z^{-n_0}X(z)$)

$$\sum_{k=0}^{N} a_k z^{-k} Y(z) = \sum_{l=0}^{M} b_l z^{-l} X(z) \tag{7.41}$$

Since the discrete transfer function $T(z)$ is defined as the z-transform of the output $y(z)$ divided by the z-transform of the input $X(z)$ (see (7.32)) we obtain from (7.41)

$$\begin{aligned} T(z) = \frac{Y(z)}{X(z)} &= \frac{\sum_{l=0}^{M} b_l z^{-l}}{\sum_{k=0}^{N} a_k z^{-k}} = \left(\frac{b_0}{a_0}\right) \frac{\prod_{l=1}^{M} (1 - c_l z^{-1})}{\prod_{k=1}^{N} (1 - d_k z^{-1})} \\ &= z^{N-M} \cdot \left(\frac{b_0}{a_0}\right) \frac{\prod_{l=1}^{M} (z - c_l)}{\prod_{k=1}^{N} (z - d_k)} \end{aligned} \tag{7.42}$$

which is a rational function in $z$. $c_l$ are the non-trivial zeros and $d_k$ are the non-trivial poles of $T(z)$. In addition $T(z)$ has $|N - M|$ zeros (if $N > M$) or poles (if $N < M$) at the origin (trivial poles and zeros). Poles or zeros may also occur at $z = \infty$ if $T(\infty) = \infty$ or $0$, respectively. Keeping in mind the correspondence between a time shift by $k$ samples and multiplication with $z^{-k}$ of the corresponding z-transform (shifting theorem), we can see that the coefficients of the polynomials in the transfer function $T(z)$ in (7.42) are identical to the coefficients of the difference equation in (7.40). Hence, we can construct one from the other.

### 7.8 The z-transform and the Discrete Fourier Transform (DFT)

When comparing the Fourier spectra of continuous-time signals of infinite duration (CTFT) and discrete-time signals of finite duration (DFT), we have already noticed that the CTFT is defined in a continuous-frequency domain while the DFT is defined for discrete frequencies. The transition from continuous-time to discrete-time by sampling corresponded to the transition from aperiodic to periodic spectra, while the transition from aperiodic signals to periodic signals - or the periodic extension $\tilde{x}[n]$ of a finite signal $x[n]$ as in the case of the DFT - corresponded to the transition from continuous spectra to discrete spectra (cf. Fig. 7.2). We have also seen above that the discrete-time Fourier transform (DTFT) for infinite sequences equals the z-transform evaluated on the unit circle for continuous frequencies. Therefore, it should come as no surprise that the DFT for

$x[n]$ samples the z-transform of $\tilde{x}[n]$ at discrete frequencies, namely at N equally spaced points on the unit circle as schematically displayed in Fig. 7.6.

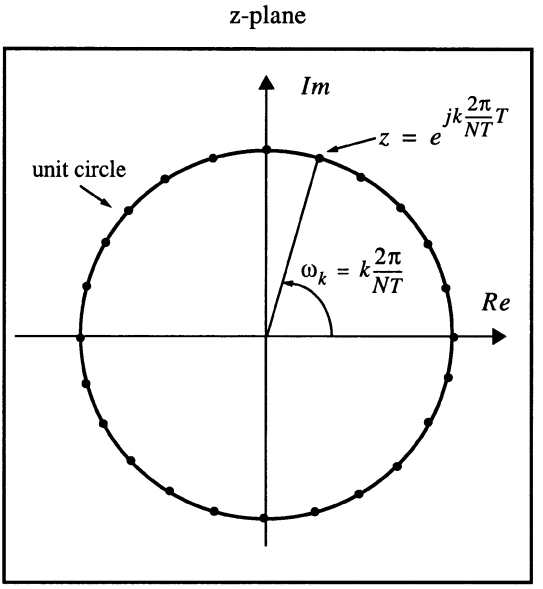



**Fig. 7.6** DFT representation in the z-plane.

## 7.9 The convolution of sequences

The concept of the discrete transfer function $T(z)$ defined in equation (7.32) enables us to calculate the z-transform of the output of a discrete filter by multiplying the z-transform of the discrete filter input signal with $T(z)$ . On the other hand, the convolution theorem for the z-transform (equation (7.36)) tells us that this is equivalent to the discrete convolution of two sequences, namely the discrete filter input signal and the discrete impulse response of the filter. For continuous-time signals we could also obtain a filter output signal by multiplying the Fourier transform of the input signal with the frequency response function of the filter and subsequent inverse Fourier transform. By analogy we could therefore expect that this is true also for sequences of finite duration. In other words, that convolution could be done in the frequency domain by simply multiplying the corresponding discrete DFT spectra. This is true, however, only if we take special precautions. Otherwise the results may be not what we expect as the following example demonstrates.

First, let us create the impulse response functions of two filters which perform just a simple time shift. The corresponding impulse response functions simply consist of time shifted impulses. For a sampling rate of 100 Hz and a time shift of 5.5 sec, the impulse response function would consist of a spike at sample 550. For a 6 sec delay filter, the spike would have to be located at sample 600. The convolution of the two filters should result in a time shifting filter which delays the input signal by 11.5 sec. Fig. 7.7, shows the results of the naive application of the convolution theorem simply multiplying the DFT spectra of the individual finite discrete signals to obtain the DFT spectrum of the output signal. For the calculation of the FFT, 1024 points were used.

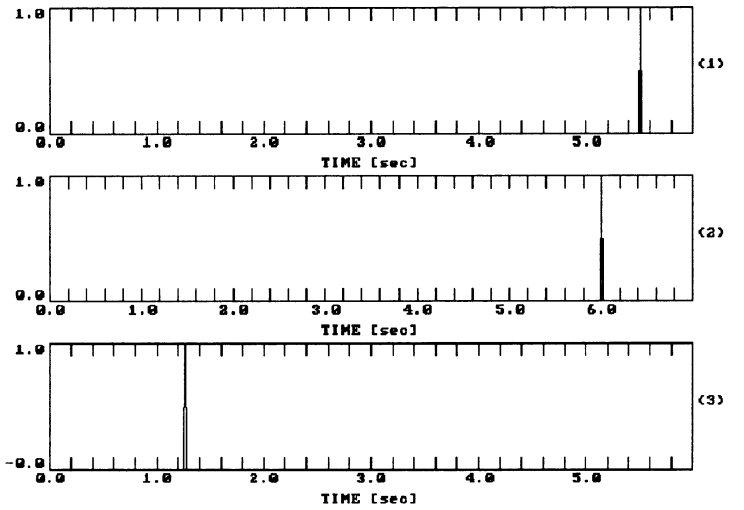



**Fig. 7.7** The wrap around effect with discrete convolution in the frequency domain. Channel 3 is the inverse DFT of the product of the DFT spectra of channels 1 and 2. The no. of points used for calculating the DFT was equivalent to 10.24 sec.

Extrapolating our experience from the convolution of continuous-time signals, we would expect that the output signal would have a spike at 11.5 sec (index 1150), which is more than we can display with 1024 points. Hence the output signal should be zero. Instead, we see that we have a spike at position 126. Well, we naively assumed that multiplying two DFT spectra gives the same result as multiplying two Fourier spectra. Instead, what we see is called *wrap around effect* or *temporal aliasing*. This is due to the fact that the DFT is essentially the Fourier series representation (equation (7.22)) of an infinite periodic sequence $\tilde{x}[nT]$ which is the periodic extension of the given finite sequence $x[nT]$. Consequently, multiplying two DFT spectra corresponds to the convolution of two infinite periodic sequences.

• *Convolution theorem for periodic sequences (circular convolution)* — If a sequence $\tilde{x}_1[n]$ is periodic with a period N, then its discrete convolution with another signal $\tilde{x}_2[m]$ of duration N is identical to the inverse DFT of the product of the corresponding DFT spectra.

Hence, naively multiplying two DFT spectra in order to obtain the same result as expected from the properties of convolution of continuous signals will in general fail to produce the desired result. For N number of points in the spectra, the result will be the same as if the input sequence were periodic with the period of N. What we observe as wrap around effect is this periodicity. Likewise, other properties of the DFT can be intuitively derived from the corresponding properties of the Fourier transform, only if we do not think in terms of the finite sequence $x[n]$ but always in terms of $\tilde{x}[n]$ which contains infinitely many replicas of $x[n]$ .

To avoid the wrap around effect we can use a simple trick called *zero padding*. If we artificially increase the period (N) by padding the signal with trailing zeros to make it larger than the largest time lag for which the input signal could possibly be affected by the impulse response, wrap around will not affect the output.

*Problem 7.4* How many samples long should the zero padding be in the example above in order to eliminate the wrap around effect?

CHAPTER 8

# The digital anti-alias filter

High performance seismic recording systems that make use of oversampling and decimation techniques reduce the requirements on the analog anti-alias filters because most of the filtering is performed in the digital domain (cf. Fig. 6.19). The quantization noise level of such systems decreases within the passband of interest with the ratio of the original Nyquist frequency to the final Nyquist frequency (oversampling ratio). On the other hand, it is desirable to keep the usable frequency range for the signal (passband of the anti-alias filter) as wide as possible. For this reason, digital anti-alias filters with very steep transition bands are needed to obtain the best resolution for a given frequency band of interest. The filter should also leave a band-limited signal which falls completely within the passband as unaffected as possible, causing neither amplitude nor phase distortions.

In order to implement digital filters, we can proceed in different ways. In terms of the concepts we have encountered so far with general LTI systems, we have seen that a filter process can be described as the multiplication of a 'spectrum' with a 'transfer function'. Since we are now dealing only with discrete filters, 'spectrum' in the most general sense means the z-transform of the discrete signal and 'transfer function' refers to the z-transform of the discrete impulse response. For filtering we multiply the discrete Fourier spectrum of the signal (DFT) with the discrete frequency response function of the filter. Of course, we should consider the wrap around effects described above. On the other hand, we can perform the filtering process in the time domain by directly evaluating the convolution sum. We can then make use of the fact that the discrete systems we are dealing with can all be described by rational transfer functions in z (7.42)

$$H(z) = \frac{\sum_{l=0}^{M} b_l z^{-l}}{\sum_{k=0}^{N} a_k z^{-k}} = \left(\frac{b_0}{a_0}\right) \frac{\prod_{l=1}^{M} (1 - c_l z^{-1})}{\prod_{k=1}^{N} (1 - d_k z^{-1})} . \tag{8.1}$$

The corresponding linear difference equation (7.40):

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{l=0}^{M} b_l x[n-l] \tag{8.2}$$

relates the output sequence $y[n]$ to the input sequence $x[n]$ . This concept naturally leads to two important classes of filters, *recursive* and *non-recursive filters*. For recursive filters, the filter output $y[n]$, at a given sample index *n*, not only depends on the input values $x[n]$ at sample index *n* and earlier samples *n-l*, but also on the output values $y[n-k]$ at earlier samples *n-k*. Here $k$ and $l$ vary from 0 to $N$ and from 0 to $M$, respectively. Hence with recursive filters, it is possible to create filters with infinitely long impulse responses (IIR filters = Infinite Impulse Response filters). It should be noted however, that not all recursive filters are IIR filters (cf. Hamming, 1977; example on p. 208). If $a_0 = 1$ and $a_k = 0$ for $k \geq 1$, $y[n]$ will depend only on the input values $x[n]$ at sample *n* and earlier samples, *n - l* (for $l$ running from 0 to *M*) and the impulse response will be finite. Such filters are called FIR (Finite Impulse Response) filters. From equations (8.1) and (8.2), it can be seen that the transfer function $H(z)$ of FIR filters is completely described by zeros ($c_l$) (except for trivial poles at the origin cf. also equation (7.42)) while that of IIR filters contains both non-trivial poles and zeros ($d_k$ and $c_l$).

The design of a filter to given specifications can be done directly by placing poles and zeros in the complex z-plane and exploiting what we know about system properties and the location of poles and zeros of the corresponding transfer function (cf. chapter 7.6). For many situations, continuous filters with desirable properties may already be known or can easily be designed. For these cases, techniques are available to map continuous filters directly into discrete ones. Some of these methods are based on sampling the impulse response of a given filter under certain constraints (e.g. impulse invariance or step invariance). Others, such as the bilinear transform, directly map the s-plane onto the z-plane. Some of the advantages and drawbacks of the individual approaches are discussed in chapter 9.2. A general treatment of these topics can be found in Oppenheim and Schafer (1989).

For modern seismic acquisition systems the kind of filter which is required for the digital anti-alias filter is most easily implemented as a *FIR filter* as is discussed below. Although a complete discussion of the problem is beyond the scope of this text, some of the advantages and disadvantages of FIR and IIR filters are listed below.

- FIR filters: They are always stable. For steep filters they generally need many coefficients, although part of this problem can be overcome by special design procedures. Both causal and noncausal filters can be implemented. Filters with given specifications such as linear phase or zero phase are easy to implement.
- IIR filters: They are potentially unstable and subject to quantization errors (Since steep filters require poles and zeros to be located close to but inside the unit circle, roundoff errors and the finite word length of the computer may cause them to wander out of the unit circle and result in potential instabilities). Steep filters can easily be implemented with a few coefficients. As a consequence, filtering is fast. Filters with given specifications such as zero phase are in general, difficult, if not impossible, to implement *exactly(!)*. In order to produce zero phase responses from recursive filters, one can filter the signal twice in opposite directions. This cancels the phase response. Obviously this can only be done with finite sequences, hence the name IIR filter is no longer appropriate Also, this approach can not be implemented as a single recursive filter in real time.

The digital anti-alias filters most commonly implemented in modern seismic recording systems belong to the class of generalized *linear phase* filters which cause no phase distortions of the signal. In general, they produce a constant time shift which can be accounted for. If the time shift is zero or is corrected for, the filter would be called a *zero phase* filter. To my knowledge, the reasons for using FIR filters in this context are primarily based on the fact that well-established design procedures exist to match given design specifications with great accuracy (McClellan et al., 1979), that very steep stable filters are needed, (these would be potentially unstable as IIR filters), that the linear phase property can be implemented exactly only as FIR filters. Most importantly, IIR filters always cause phase distortion **within the passband** of the filter.

Fig. 8.1 shows an example for a typical FIR filter impulse response used in modern seismic recording systems.

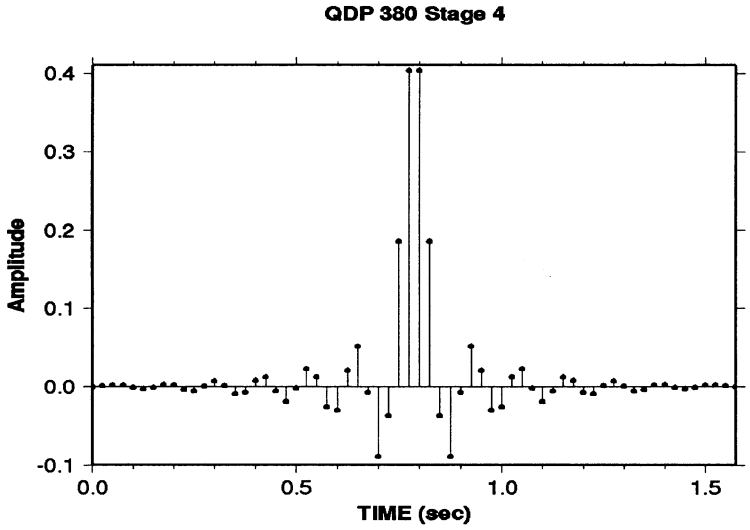



**Fig. 8.1** FIR filter impulse response of the stage 4 of the QDP 380 digitizer by Quanterra, Inc. (J. Steim, pers. comm. 1993). It is used as digital anti-alias filter in the 40 Hz data stream while it is decimated to 20 Hz.

Such filters with symmetrical impulse responses are often called two-sided or acausal for reasons discussed below. The impulse response oscillates with frequencies close to the corner frequency of the filter, an effect called Gibb's phenomenon (Oppenheim and Schafer, 1989). It could be reduced at the cost of the width of the filter pass band by decreasing the steepness of the filter. This, however, would also reduce the usable frequency range for the signal.

As a consequence of the symmetry of the filter impulse response, the onsets of very impulsive signals may be obscured by precursory ('acausal') oscillations and become hard, if not impossible, to determine. In Fig. 8.2, an example for FIR filter generated onset distortion is shown for a record obtained at the station HUIG of the National Mexican Network. The trace shows the P-wave portion on the vertical component sampled at 20 Hz. Before decimation from 40 to 20 Hz it has been filtered with the FIR filter response shown in Fig. 8.1. As can be seen from Fig. 8.2, the use of a symmetric FIR filter in this case obscures the seismic onset considerably and generates spurious acausality.

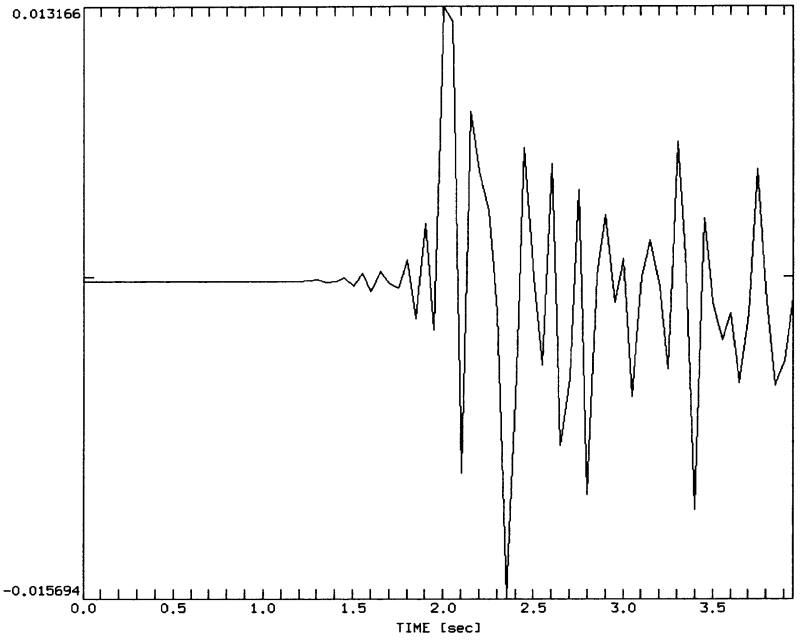



**Fig. 8.2** Vertical component of a P wave record with FIR filter generated precursory signals obtained at the station HUIG of the National Mexican Network (data provided by M.-P. Bouin, pers. comm., 1995). The sampling rate is 20 Hz. The digital anti-alias filter used before decimation from 40 to 20 Hz is shown in Fig. 8.1.

One way to deal with this problem which is discussed below, is to construct a filter which removes the acausal (left-sided) component of the filtered signal and replaces it by its equivalent causal (right-sided) component. Besides its own merits, this particular approach provides a good opportunity to apply some of the concepts we have been studying so far.

In this context, the digital seismogram **before decimation** is viewed in terms of the z-transform as

$$\tilde{Y}(z) = F(z) \cdot z^{lp} \cdot \tilde{X}(z) \tag{8.3}$$

Here, $\tilde{Y}(z)$ represents the z-transform of $\tilde{y}[n]$, the digitally low-pass filtered seismic trace before decimation. $F(z)$ is the z-transform of the FIR filter and $\tilde{X}(z)$ is the z-transform of the input signal $\tilde{x}[n]$. Since the FIR filter described by $F(z)$ is a linear phase filter which produces a constant time shift of $lp$ samples[1], the filtered signal will

1. Note that $lp$ may be real, it does not have to be an integer number.

be delayed by $lp$ samples. This delay is corrected by shifting the filtered trace back in time by the same amount. In terms of the z-transform this corresponds to multiplication with $z^{lp}$. Hence $F(z) \cdot z^{lp}$ corresponds to a zero phase filter in which the linear phase component of $F(z)$ is corrected. Since the spurious acausalities are generated by the symmetry of the linear phase FIR filter, we will focus on the treatment of $F(z)$ and will consider the time shift $z^{lp}$ separately.

## 8.1 Removing the maximum phase component of a FIR filter

The method discussed below is based on the fact that the impulse response of the linear phase FIR filter can be separated in its minimum and its maximum phase component. The left-sided or 'acausal' part is caused by the maximum phase component with the transfer function $F_{max}(z)$ while the right-sided or 'causal' part corresponds to the minimum phase component with the transfer function $F_{min}(z)$. The complete filter transfer function $F(z)$ is the product of $F_{min}(z)$ and $F_{max}(z)$. Our goal is to remove the maximum phase component of $F(z)$ from the filter output trace without changing the amplitude characteristics of the filter process. In particular we want to replace $F_{max}(z)$ by its minimum phase equivalent, which will be denoted $\mathcal{MinPhase}\{F_{max}(z)\}$. In other words we want to create an allpass filter similar to the one discussed in Problem 3.4. Several methods to attack this problem practically have been reported (Fowler, 1992; Wielandt, 1992; see also Quanterra technical note on linear and minimum-phase filters, Quanterra Inc., Harvard, MA, 1992).

Here we will discuss the construction of a recursive correction filter similar to the approach of Wielandt (1992). In contrast to the method of Wielandt (1992), however, we are going to use the representation of $F(z)$ by its singularities to determine the maximum phase component $F_{max}(z)$. Since we are dealing with FIR filters, we know from chapter 7.6 and from equation (8.1) that the transfer function $F(z)$ can be described completely by its zeros. For the present context we can ignore the fact that this is only true if we ignore $z = 0$ and $\infty$ (cf. Problem 7.3). The zeros inside the unit circle ($c_i^{min}$) correspond to the minimum phase component $F_{min}(z)$, while the zeros outside the unit circle ($c_i^{max}$) correspond to $F_{max}(z)$. Therefore, we can construct the minimum and maximum phase components of $F(z)$ by considering the contributions from zeros inside and outside of the unit circle separately.

What we need to consider next is the determination of $\mathcal{MinPhase}\{F_{max}(z)\}$. For this purpose we will use the fact that replacing the position of a singularity by its complex conjugate reciprocal does not change the amplitude part of the corresponding transfer function except for a constant scaling factor (the reciprocal of the distance of the original root to the origin). We have encountered the equivalent property for continuous systems in the context of Problem 3.1 where we changed poles in the s-plane symmetrically to the imaginary axis. By taking the reciprocal in the z-plane we change the position of singularities symmetrically to the unit circle. This demonstrates again that with a few simple rules, such as replacing the role of the imaginary axis in the s-plane by the role of the unit circle in the z-plane, important properties of discrete systems can intuitively be understood from similar properties of continuous systems.

Regarding $F_{max}(z)$, if all zeros move inside the unit circle the transfer function becomes minimum phase. Therefore, replacing all zeros lying outside the unit circle ($c_i^{max}$) by their complex conjugate reciprocal values ($1/c_i^{max*}$) and multiplying the resulting transfer function with the product of all distances of the original roots from the origin yields a minimum phase transfer function with the same absolute value as $F_{max}(z)$.This, however, is exactly the desired result $\mathcal{MinPhase}\{F_{max}(z)\}$.

Practically, the transformation is done by replacing $F_{max}(z)$ by $F_{max}(1/z)$. This allows us to write

$$\mathcal{MinPhase}\{F_{max}(z)\} = F_{max}(1/z) \; . \tag{8.4}$$

According to the properties of the z-transform discussed in chapter 7.6, however, this simply corresponds to inverting the impulse response of $F_{max}(z)$ in time. In other words, to make the response of $F(z)$ causal, one deconvolves the maximum phase component and then convolves the remaining sequence with the time-inverted maximum phase component.

In terms of the z-transform, the corrected seismogram $Y(z)$ can be expressed by division and multiplication

$$Y(z) = \frac{1}{F_{max}(z)} \cdot F_{max}(1/z) \cdot \tilde{Y}(z) \tag{8.5}$$

This creates a problem! Since $F_{max}(z)$ has only zeros outside the unit circle, $1/F_{max}(z)$ will have only poles outside the unit circle, for which only the left-sided

impulse response will be stable (cf. Fig. 7.5). However, if we invert all signals in time before filtering

$$Y(1/z) = \frac{1}{F_{max}(1/z)} \cdot F_{max}(z) \cdot \tilde{Y}(1/z) \tag{8.6}$$

the impulse response corresponding to $1/F_{max}(1/z)$ becomes a stable causal sequence in nominal time. Hence, in nominal time, the deconvolution of the maximum phase component $F_{max}(z)$ poses no stability problems.

### 8.1.1 THE DIFFERENCE EQUATION

For the practical implementation we rewrite equation (8.6) to obtain

$$F_{max}(1/z) \cdot Y(1/z) = F_{max}(z) \cdot \tilde{Y}(1/z) \tag{8.7}$$

This can be written as

$$A'(z) \cdot Y'(z) = B'(z) \cdot X'(z) \tag{8.8}$$

Here, $Y'(z)$ and $X'(z)$ correspond to $Y(1/z)$ and $\tilde{Y}(1/z)$, while $A'(z)$ and $B'(z)$ correspond to $F_{max}(1/z)$ and $F_{max}(z)$, respectively. Written as convolution sum, this becomes

$$\sum_{k=-\infty}^{\infty} a'[k] \cdot y'[i-k] = \sum_{l=-\infty}^{\infty} b'[l] \cdot x'[i-l] \tag{8.9}$$

If we assume $F(z)$ contains $mx$ zeros outside the unit circle, the wavelets $a'[k]$ and $b'[l]$ will be of length $mx + 1$. Hence, equation (8.9) becomes

$$\sum_{k=0}^{mx} a'[k] \cdot y'[i-k] = \sum_{l=0}^{mx} b'[l] \cdot x'[i-l] \tag{8.10}$$

or

$$y'[i] \cdot a'[0] + \sum_{k=1}^{mx} a'[k] \cdot y'[i-k] = \sum_{l=0}^{mx} b'[l] \cdot x'[i-l] \tag{8.11}$$

which can also be written

$$y'[i] = -\sum_{k=1}^{mx} \frac{a'[k]}{a'[0]} \cdot y'[i-k] + \sum_{l=0}^{mx} \frac{b'[l]}{a'[0]} \cdot x'[i-l] \tag{8.12}$$

This is formally identical to equation (8.2). With

$$a[k] = -\frac{a'[k]}{a'[0]} = \frac{f_{max}[mx-k]}{f_{max}[mx]} \quad \text{for } k = 1 \text{ to } mx \tag{8.13}$$

and

$$b[l] = \frac{b'[l]}{a'[0]} = \frac{f_{max}[l]}{f_{max}[mx]} \quad \text{for } l = 0 \text{ to } mx \tag{8.14}$$

equation (8.9) becomes

$$y'[i] = \sum_{k=1}^{mx} a[k] \cdot y'[i-k] + \sum_{l=0}^{mx} b[l] \cdot x'[i-l] \tag{8.15}$$

Hence, if we know the $mx+1$ coefficients of the maximum phase portion of a linear phase FIR filter, we can employ equation (8.15) to calculate $y'[i]$, the time reversed sequence for which the non-causal part of the FIR filter has been replaced by its equivalent minimum phase part. After filtering, we can obtain $y[i]$ by simply flipping $y'[i]$ back in time. Finally (cf. equation (8.5)), $Y(z)$ can be written as

$$Y(z) = \tilde{F}(z) \cdot \tilde{Y}(z) \tag{8.16}$$

$\tilde{F}(z)$ is a time advanced minimum phase filter advancing the output signal by $lp$ samples. To account for this time shift in the corrected seismogram, either the time tag of the trace needs to be changed or the trace needs to shifted physically by $lp$ samples.

### 8.1.2 PRACTICAL CONSIDERATIONS

So far, we have not been concerned with the details of practical implementation. Here, we will make an exception, since this will allow us to discuss some problems commonly arising in this context.

Prerequisite for all attempts to eliminate the noncausal part of a FIR filter response is the determination of its maximum phase component. As long as the number of coefficients of the FIR filter impulse response is not too large, we can exploit the position of zeros of $F(z)$ for this purpose. Practically this can be done by using a polynomial rooting technique on the impulse response of the FIR filter. This is justified by the fact that the coefficients of the impulse response of the FIR filter are also the coefficients of its z-transform $F(z)$ expressed as polynomial (cf. equation (7.31)). Once all the roots are determined, those which lie outside the unit circle ($c_i^{max}$, $i = 1$ to $mx$) are collected. Subsequently, we can rewrite $F_{max}(z)$ as

$$F_{max}(z) = f[m] \prod_{i=1}^{mx} (z - c_i^{max}) \tag{8.17}$$

with $f[m]$ the last coefficient of the impulse response of the FIR filter. The desired corresponding maximum phase impulse response $f_{max}[n]$ can be obtained by evaluating (8.17) numerically. It is convenient to do this on the unit circle because, as we have seen before, this is equivalent to calculating the inverse DFT for $F_{max}(e^{j\omega T})$. In other words, eqn. (8.17) evaluated for $z = e^{j\omega T}$ provides the spectral values of the DFT of $f_{max}[n]$. The values of $f_{max}[n]$ can then directly be plugged into the difference equation (8.15) and used for filtering.

The advantage of the rooting approach is that it yields both $F_{max}(z)$ as well as $mx$, the length of the maximum phase portion of an arbitrary asymmetric FIR filter. For FIR filter responses with less than 100 coefficients commonly available techniques such as implemented in Mathematica (Wolfram, 1991) or described by Press et al., (1992) have been found to work sufficiently well. For longer filters, other techniques such as cepstrum analysis (Quatierie and Tribolet, 1979; Oppenheim and Schafer, 1989) can be used to determine the maximum phase component (e.g. Wielandt, 1992). The discussion of these methods, however, is beyond the scope of this text.

With the numerical implementation of the calculation of $f_{max}[n]$ we face two common, sometimes confusing problems:

First, rooting algorithms often provide the roots of polynomials $P(x)$ with $P(x) = \sum_{i=0}^{m} a_i x^i$ (e.g. Press et al., 1992). In contrast, the definition of the z-transform

which we have been using so far (eqn. (7.31)), $F(z) = \sum_{k=0}^{m} f[k] \cdot z^{-k}$ contains only negative powers of $k$.

Secondly, the most common implementations of the DFT (e.g. Press et al., 1992) calculates $\tilde{X}[k] = \sum_{n=0}^{N-1} \tilde{x}[n] e^{j2\pi kn/N}$ while we have defined the DFT using a negative exponent in equation (7.24).

We could deal with the first problem by simply multiplying $F(z)$ by $z^m$ which would correspond to a negative time shift of $f[n]$ by $m$ samples and would introduce additional zeros at the origin (see Problem 7.3). Because of the symmetry between DFT and inverse DFT, we could deal with the second problem by using the code for the DFT from Press et al. (1992) to calculate - after appropriate scaling- the inverse DFT according to the definition we are used to. However, there is an easier way out which solves both problems simultaneously, if we temporarily use a different definition of the z transform

$$\underset{\sim}{X}(z) \equiv \sum_{n=-\infty}^{\infty} x[n] \cdot z^{n} \tag{8.18}$$

With this definition, we can directly use $x[n]$ as input for the rooting algorithm. Furthermore, the evaluation of the z-transform $\underset{\sim}{X}(z)$ on the unit circle directly corresponds to the evaluation of the inverse DFT in the formalism of Press et al. (1992). What is different, however, are the relationships between the phase properties and the locations of the poles and zeros. Using the definition of equation (8.18), minimum phase filters must have all their poles and zeros outside the unit circle while maximum phase filters have their singularities located inside the unit circle.

Below, the performance of the rooting approach for the FIR filter response of Fig. 8.1 is demonstrated. Fig. 8.3 shows the distribution of roots in the z-plane using the convention of equation (8.18).

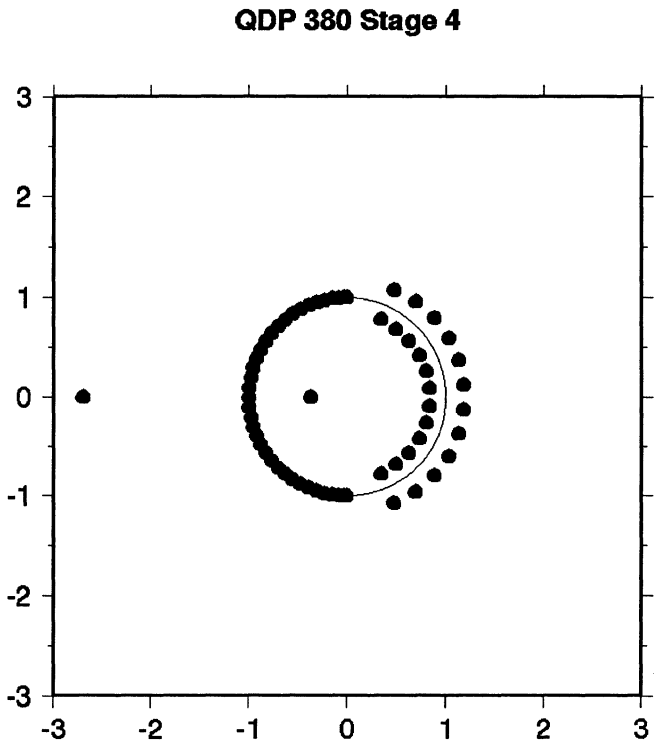



**Fig. 8.3** Distribution of roots for the FIR filter response shown in Fig. 8.1 (stage 4 of the QDP 380 digitizer by Quanterra, Inc.).

We can notice that most of the roots in the left half z-plane fall directly onto the unit circle. From the mapping of the s-plane to the z-plane (cf. Fig. 7.3) we know that these correspond to the upper half of the frequency band. They generate the stop band of the filter by suppressing all energy in that frequency range. In the low frequency range (closer to the positive real axis) the roots are located away from the unit circle. In the notation of equation (8.18) the roots inside the unit circle correspond to the maximum phase component while the roots outside the unit circle correspond to the minimum phase component of the FIR filter. Based on the position of the roots we can decompose the FIR filter response of Fig. 8.1 into its minimum and maximum phase component and the component corresponding to the roots directly on the unit circle (Fig. 8.1).

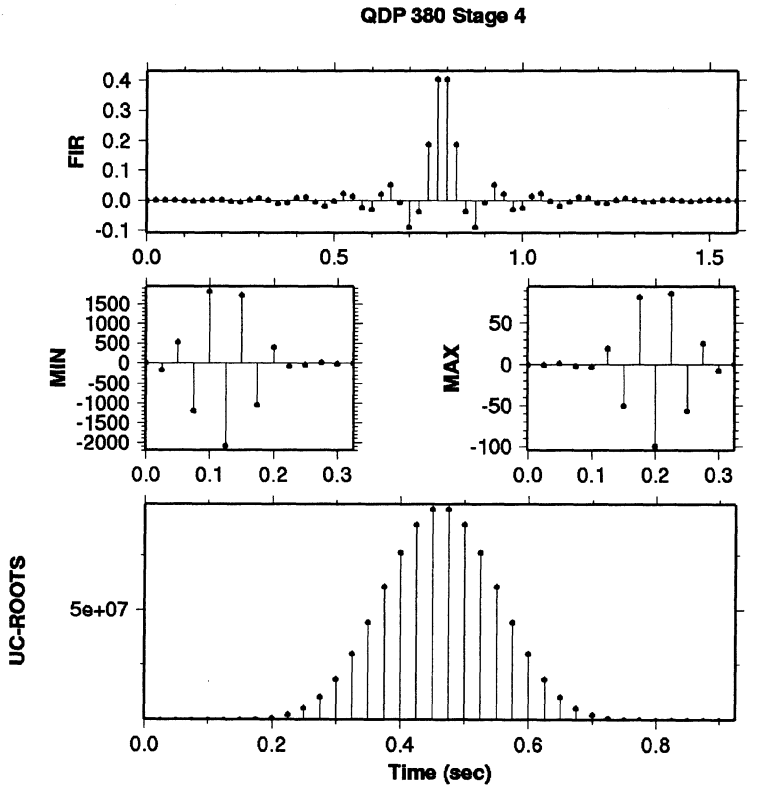



**Fig. 8.4** Decomposition of the FIR filter shown in Fig. 8.1 into purely minimum phase (center left), purely maximum phase (center right) and the component corresponding to roots directly on the unit circle (bottom panel).

In terms of the phase properties, roots on the unit circle play a special role. In some respect they can be treated as belonging to the minimum phase component while in others they cannot (Oppenheim and Schafer, 1989). For example, in contrast to purely minimum phase signals, their inverse is not minimum phase. Although their signal component is symmetrical, in the context of removing the acausal filter response from seismic records they can be ignored completely. This is demonstrated in the following problem.

*Problem 8.1* Do roots on the unit circles belong to the minimum or the maximum phase part of the impulse response? Start by considering a short 3 point wavelet with roots exactly on the unit circle. Why can unit circle roots be ignored for the purpose of removing the acausal filter response from seismic records?

To replace the left-sided part of the FIR filter impulse response shown in Fig. 8.1 by its minimum phase equivalent, we can use the coefficients of its purely maximum phase component shown in Fig. 8.1 as input for the correction filter difference equation (8.15). The result is shown in Fig. 8.5.

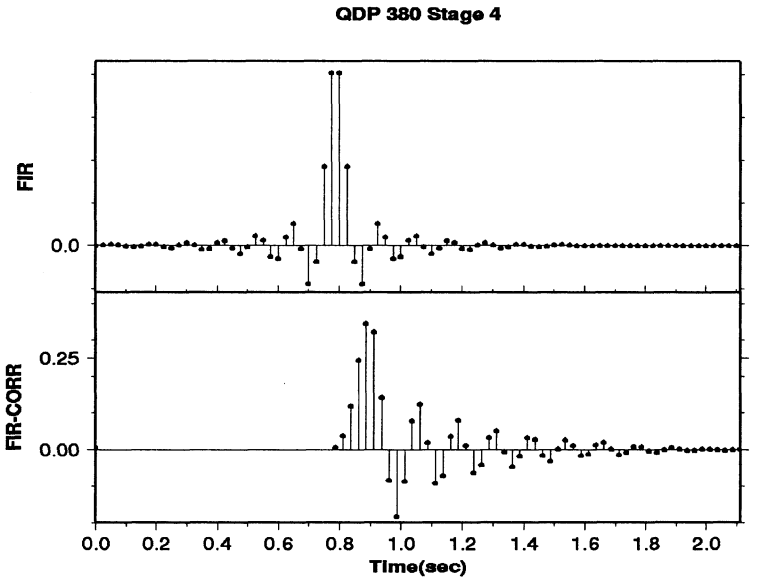



**Fig. 8.5** Replacing the maximum phase component of the FIR filter response shown in Fig. 8.1 by its minimum phase equivalent. The top trace shows the original FIR filter response (with zeros appended). The bottom trace shows the result of the correction. In addition, a time shift to correct for the linear phase component of the FIR filter response has been applied to the bottom trace.

To apply the same correction filter to the data trace shown in Fig. 8.2, the data were first interpolated and resampled at 40 Hz to obtain the same sampling frequency as for the original FIR filter response. This was done in the frequency domain by appending zeros to the DFT followed by an inverse DFT (cf. Proakis and Manolakis, 1992). Subsequently, the correction filter was applied. The resulting traces are displayed in Fig. 8.6. Notice the absence of the precursory signals before the P wave onset which dramatically facilitates the determination of the P wave onset time. However, it should be noted that the resulting waveforms are now strongly modified by the phase response of the correction filter.

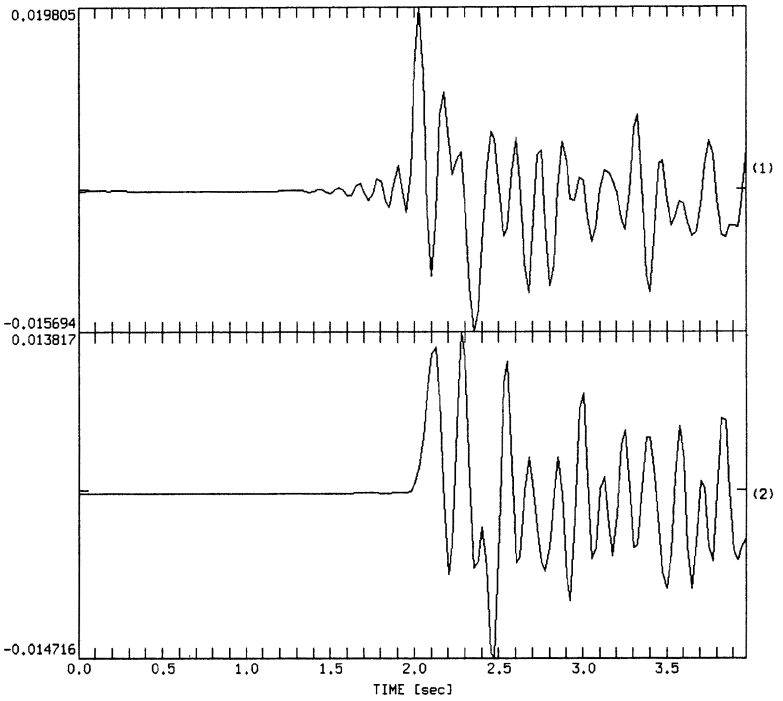



**Fig. 8.6** Removing the maximum phase component of the FIR filter response from the signal shown in Fig. 8.2. The top trace shows the signal of Fig. 8.2 interpolated and resampled at 40 Hz. The bottom trace shows the corrected signal.

Besides in the context of onset time determination, precursory signals caused by the FIR filter also need to be considered in another context. They may severely affect the identification of source related rupture nucleation phases which have recently been suggested to exist based on the analysis of near-source observations of earthquakes of different magnitudes (Ellsworth and Beroza, 1995; Iio, 1995). These phases have been observed on analog and analog filtered digital recordings as intervals of weak motion preceding the strong motion of the main rupture. While the precursory signals in Fig. 8.2 can easily be identified as FIR filter related from their monochromatic oscillatory character, this may not always be the case. An example for such a situation is displayed in Fig. 8.7. The top trace shows the uncorrected P wave onset for a local event recorded at the station HEI of the SIL network in Southern Iceland (Stefánsson et al., 1993).

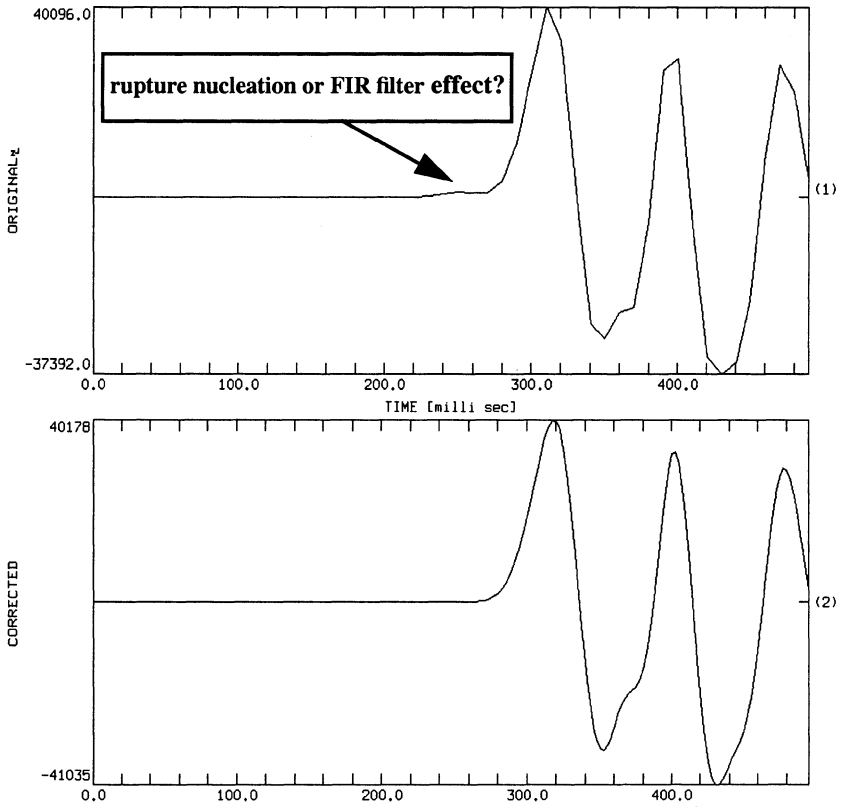



**Fig. 8.7** Vertical component of the P wave onset for a local event recorded at the station HEI of the SIL network in Southern Iceland (Stefánsson et al., 1993; data provided by S. Jakobsdottir, pers. comm., 1995). The decimated data sampled in 100 Hz are shown in the top trace. Notice the precursory signal. The bottom trace shows the corrected trace after interpolation to 400 Hz and removal of the maximum phase component of the FIR filter response of the SIL system.

Without further knowledge, the precursory signal in the top trace could easily be mistaken for a rupture nucleation phase. Only from comparison with the corrected signal it becomes clear that it is FIR filter related. Since other filter related effects may affect the analysis of seismic onsets additionally (cf. chapter *10.2 The properties of seismic onsets*), the identification of rupture nucleation phases but also of slow initital phases requires extreme care to distinguish them from instrumental effects (Scherbaum, 1997; Scherbaum and Bouin, 1997).

So far the effects of the FIR filter have been considered **before decimation**. Therefore, the removal of the maximum phase component of the filter response from a seismic trace required interpolation and resampling of the decimated filtered trace. For computational efficiency it would be very desirable to perform the FIR filter correction directly on the decimated trace. This would require the calculation of the coefficients of the correction filter from the decimated filter response. But which of the decimated filter responses should be taken? For a given decimation ratio $d$, we can obtain $d$ different decimated traces if we perform the decimation in the time domain. For the SIL system (Stefánsson et al., 1993) this is illustrated in Fig. 8.8.

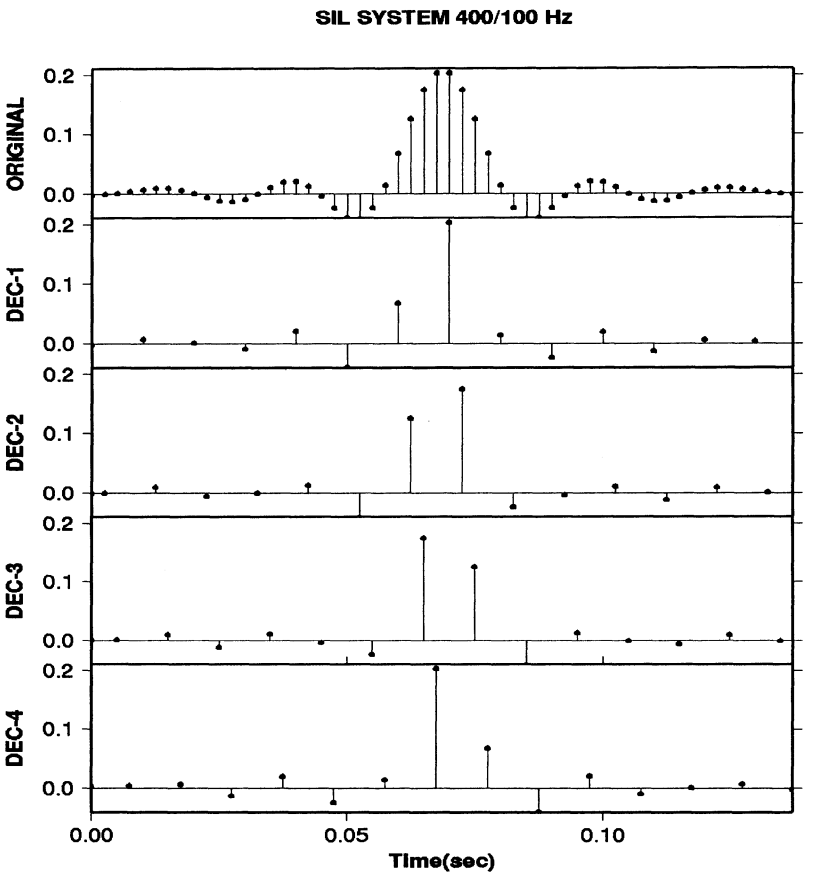



**Fig. 8.8** FIR filter impulse response for the SIL system (Stefánsson et al., 1993) for the decimation stage 400 to 100 Hz. The original response sampled at 400 Hz is shown in the top trace. Displayed below are the different possible decimated responses sampled at 100 Hz. Decimation scheme DEC-1 includes samples 1,5,9,13,...; DEC-2 samples 2,6,10,14,...; DEC-3 samples 3,7,11,15,... and DEC-4 samples 4,8,12,16,... of the original response.

Here the final decimation stage involves a decimation from 400 to 100 Hz. This decimation ratio allows 4 different decimation schemes including either the samples 1, 5, 9, 13 ,... or 2, 6, 10, 14 ,... or 3, 7, 11, 15 ,... or 4, 8, 12, 16 ,..., respectively. The phase spectra of the resulting decimated sequences in Fig. 8.8 (DEC-1 to DEC-4) differ by a linear component corresponding to the one sample time shift between the individual decimation schemes. In the decimated traces this becomes a time shift of $1/d = 0.25$ samples. This time shift directly maps into linear phase components of the corrected traces which is demonstrated in Fig. 8.9.

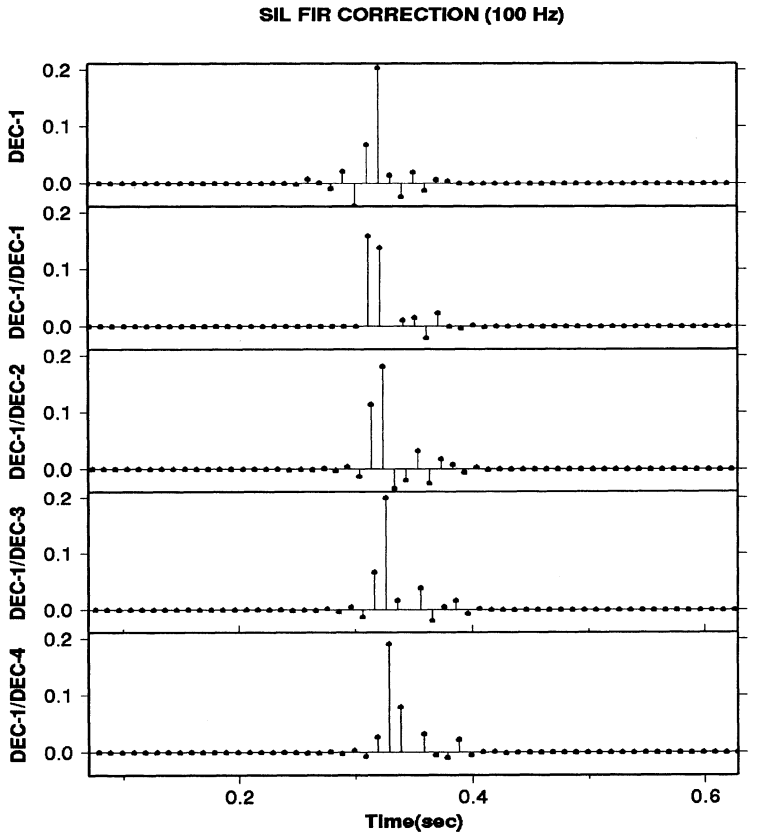



**Fig. 8.9** Removal of the maximum phase component from the filtered decimated trace. The top trace shows the decimated FIR filter response according to decimation scheme DEC-1 (cf. Fig. 8.8). This was used as 'data trace'. Traces 2 to 5 show the result of applying correction filters calculated for the four possible decimation schemes DEC-1 to DEC-4 (see label on the vertical axis).

The decimated filter response using decimation scheme DEC-1 was taken as 'data sequence' (top trace in Fig. 8.9). Correction filters were calculated from the decimated FIR filter impulse response for all four possible decimation schemes DEC-1 to DEC-4. They were subsequently applied to the top trace in Fig. 8.9. Only the correction filter for

the decimation scheme actually used (DEC-1/DEC-1, trace 2) yields the desired causal result. The other traces still show considerable precursory oscillations. These are caused by the mismatch of the actual and 'assumed' decimation scheme. The relation between precursory oscillations and time shifts for impulsive signals will be discussed in detail in chapter *10.2.4 Filter delays and onset time.* Here I leave it with the statement that they are generated by linear phase components corresponding to time shifts which are non-integer multiples of the sampling rate (for the SIL system multiples of 0.25 samples). Since these time shifts are known for the traces in Fig. 8.8 (since we know exactly which decimation schemes were used) we can correct for it. This is done by adding the corresponding linear phase components to the DFT spectra of traces 3 - 5 in Fig. 8.9 and subsequent inverse DFT. The resulting traces are displayed in Fig. 8.10.

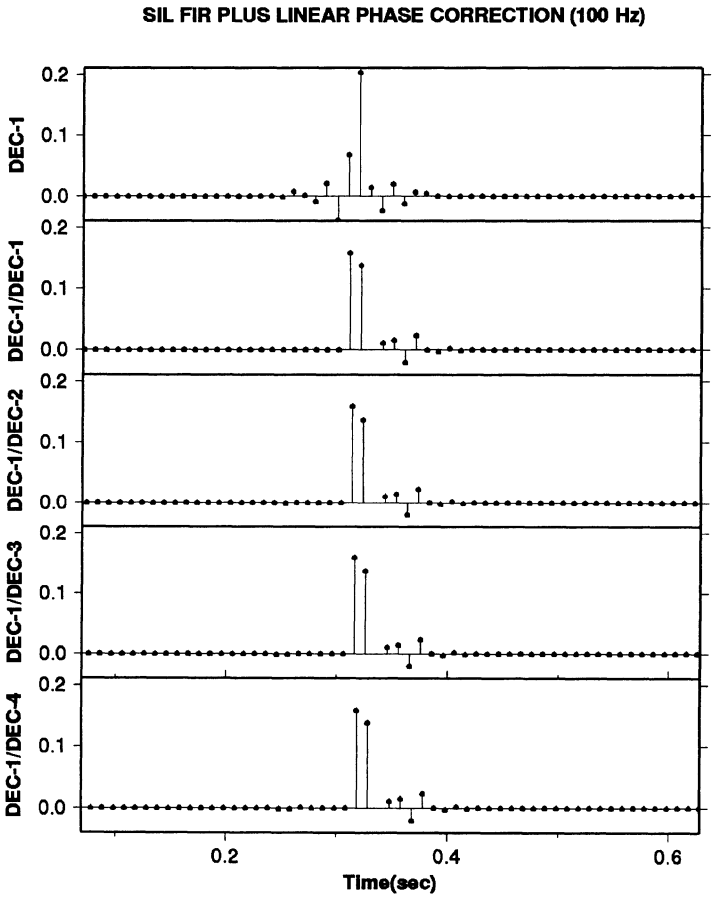



**Fig. 8.10** Correction for the linear phase component caused by the mismatch of actual and 'assumed' decimation scheme for the traces displayed in Fig. 8.9.

With this additional linear phase correction, the replacement of the maximum phase component of the decimated FIR filter response by its minimum phase equivalent performs well. With real data, however, there is no way of knowing what this linear phase component in a particular case would be. The only fact that is known is that it is a multiple of $1/d$. For a decimation ratio of $d = 4$, there is therefore a 75% chance for a decimation scheme mismatch and hence the same chance for related problems. This of course would defeat the purpose of the correction filtering in total, although, for the purpose of onset time determination the result may still yield some improvements.

For other purposes such as the identification of nucleation phases, however, the high resolution which is needed requires the interpolation and resampling of the data traces in the sampling rate before decimation.

*Problem 8.2* Demonstrate the performance of the removal of the maximum phase component of the FIR filter response with real data using DST. Data traces can be loaded into DST using the *File -> Open File with Data Trace* option from the main menu. The subdirectory *FIR20HZ* contains data examples sampled at 20 Hz with visible precursory FIR filter effects prior to the P-wave onset (e. g. file example20HZ). First, interpolate trace to 40 Hz using option *Interpolate* in the *Utilities* menu. Subsequently use option *FIR2CAUS* in the *Utilities* menu with the correction filter file quant40Hz.prt.

CHAPTER 9

# Inverse and simulation filtering of digital seismograms

We have now looked at all the essential building blocks of modern seismic recording systems (cf. Fig. 6.19) and seen how they affect recorded seismic signals. We have learned to model the individual components (seismometer, ADC, etc.) using the concepts of continuous and discrete linear time invariant systems (transfer function, frequency response function, impulse response function) and how their interactions can be described by convolution in both the time and frequency domains (convolution theorem). We have become acquainted with the elegant and powerful concept of poles and zeros and learned what they can tell us about important system properties such as minimum phase, maximum phase, etc. In addition, we have seen how we can use poles and zeros to design simple, special-purpose digital filters. After reviewing the theoretical background of the Fourier-, Laplace-, and z-transforms, we have taken a detailed look at the digital antialias filter. We have seen how we can cure the precursory ringing problem of high frequency signals and discussed the design and implementation of a recursive correction filter.

So far we have been mainly concerned with the task of filtering, that is modelling an output signal for an arbitrary input signal given a set of known system properties. In this chapter we will deal with the simulation problem - the conversion of digital (broad-band) records into those from different seismograph systems. This process is important because it allows us to determine signal amplitudes or onset times in a manner consistent with other observatories. The simulated systems will most commonly belong to the standard classes of instruments described by Willmore (1979). Since there is no single, optimum class of instruments, different instrument types must be considered for the detection and analysis of different types of seismic waves. For example, many seismologists feel that high frequency teleseismic body waves are best analysed on recordings of SP-instruments (class A), long-period body waves and teleseismic surface waves are best analysed on recordings of LP-instruments (class B), and that regional body and surface waves are best analysed from recordings of intermediate band (class C) instruments. In addition, the Wood-Anderson instrument is of special historical importance for the determination of local magnitude.

Fig. 9.1 gives, as an example, the moduli of the frequency response functions of commonly used standard instruments.

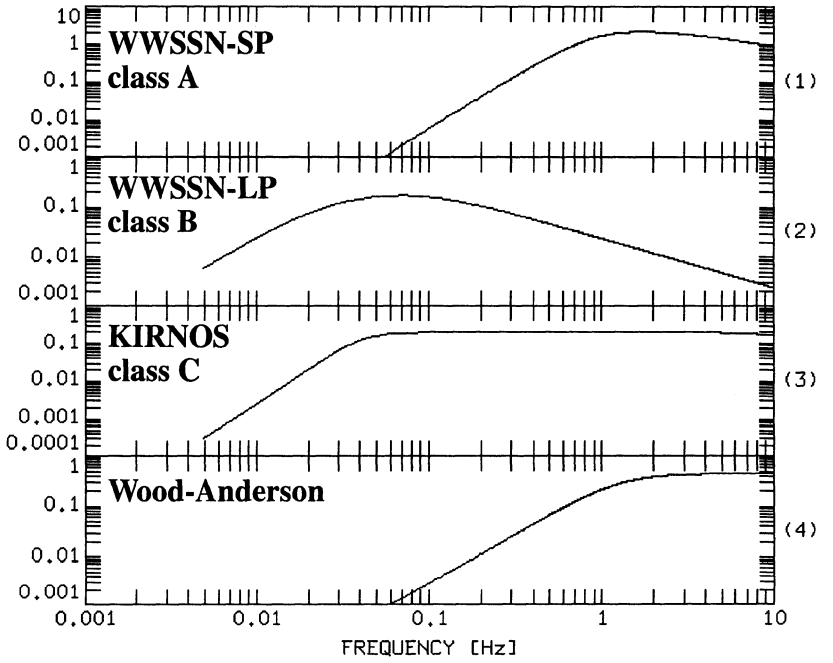



**Fig. 9.1** Shapes of the moduli of the displacement frequency response functions of commonly used standard seismograph systems.

To simulate a given seismograph system, we must first remove the effects of the original recording system. This is done by filtering the digital records of a given seismograph system with a filter that approximates the inverse response of the actual recording system. Then the resulting time series must be filtered with a response function which simulates the effects of the desired instrument. If $T_{act}(z)$ and $T_{syn}(z)$ are the transfer functions of the actual recording instrument and the instrument to be synthesized, and $Y_{act}(z)$ and $Y_{sim}(z)$ are the z- transforms of the recorded and the simulated seismograms, respectively, the simulation procedure can be described as follows:

$$Y_{sim}(z) = \frac{T_{syn}(z)}{T_{act}(z)} \cdot Y_{act}(z) = T_{sim}(z) \cdot Y_{act}(z) \quad . \tag{9.1}$$

Conceptually the simulation can be viewed as the deconvolution of the impulse response of the actual instrument from the recorded seismogram (division by $T_{act}(z)$ in (9.1)) - which corresponds to the restitution of the input ground motion signal - and a filtering problem (multiplication with $T_{syn}(z)$ in (9.1)). Here $T_{sim}(z)$ is the transfer function of a simulation filter which incorporates both these processes.The procedure is further illustrated in Fig. 9.2.

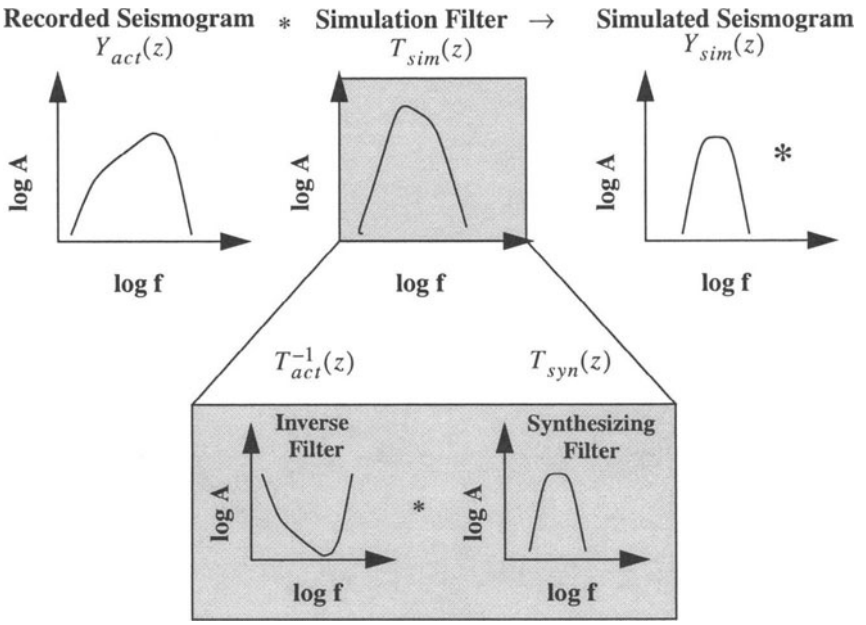



**Fig. 9.2** The simulation of digital seismographs. The simulation filter can be thought of as a combination of an inverse filter for the actual recording system and a synthesizing filter for the simulated recording system. Displayed are schematic sketches of the amplitude frequency response functions of the contributing subsystems.

## 9.1 Stability problems

Fig. 9.2 is a highly idealized plot of the real situation since both the simulated and the

actual recording system are treated as discrete systems although in reality the actual recording system contains continuous components. Furthermore, real seismograms contain spurious signal components, not related to ground motion, for which we use the generic term 'noise'. Noise can be generated in numerous ways either within the seismometer system (e.g., as thermal noise, by effects of nonlinearities, by barometric pressure effects on the sensor) or during the digitization process (limited dynamic range of the ADC, round-off errors). Here it is only important to note that noise can destabilize the simulation filter output. In quantitative terms, the signal energy at a given frequency divided by the remaining energy (noise) is called signal-to-noise ratio or SNR (Zmeskal and Plesinger, 1995).

Some consequences of noise are discussed below. To simplify we will ignore the synthesizing filter for the moment and restrict the discussion to the inverse filter process. Let us assume that we record the signal of a seismic source for which the spectrum looks like the *source spectrum* shown on the left panel in Fig. 9.3. This could correspond to the displacement spectrum of an idealized earthquake source. For the recording system sketched in the center, the recorded spectrum would look like the spectrum displayed in the rightmost panel.

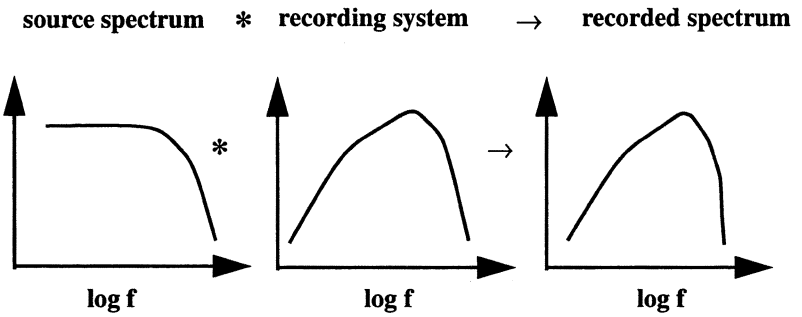



**Fig. 9.3** Recording the displacement spectrum of an idealized earthquake source.

The ideal noisefree case as shown in Fig. 9.3, the source spectrum could be recovered completely (except for zero frequency and the Nyquist frequency) by inverse filtering. This is illustrated in Fig. 9.4.

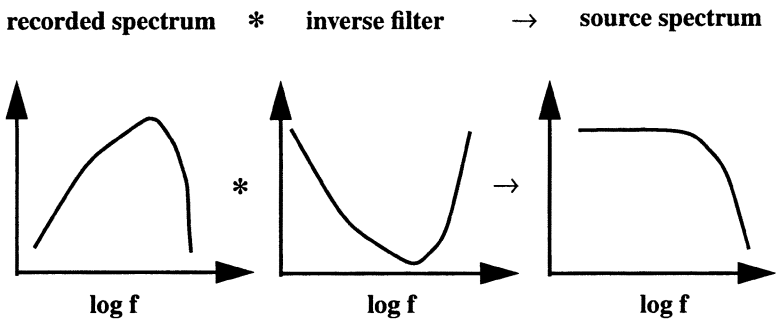



**Fig. 9.4** Recovering the source spectrum by inverse filtering in the noisefree case.

When the recorded spectrum contains noise, however, the spectrum resulting from the inverse filtering will be distorted as is shown in Fig. 9.5. The signal-to-noise ratio decreases rapidly outside the pass-band of the recording instrument (to values less than one, see Fig. 9.5 ) while in the same frequency band the magnification of the inverse filter is largest. Thus, noise in this frequency band will be amplified. If the noise component is large enough, the output of the inverse filter may become unstable. As a consequence, the instrument response can only be deconvolved within a certain *valid frequency band* (Fig. 9.5) in the presence of noise. This *valid frequency band* is not a fixed parameter but a function of the noise level at the time of the recording.

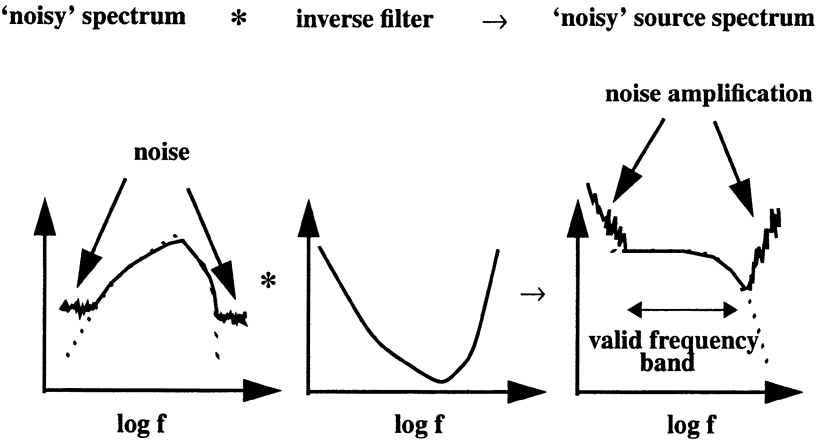



**Fig. 9.5** Noise amplification by inverse filtering. The solid line in the left panel shows the signal plus noise while the noisefree signal is shown by the dashed line.

Since the noise amplification is stronger when the slopes of the frequency response functions are steeper, the valid frequency band depends on both the signal-to noise ratio and the slope of the frequency response function of the recording systems.

The influence of noise is demonstrated below using the recording of the vertical component of an $m_b = 7.5$ earthquake in the Kurile islands recorded at station Pruhonice (PRU) of the Czech Academy of Sciences (Fig. 9.6).

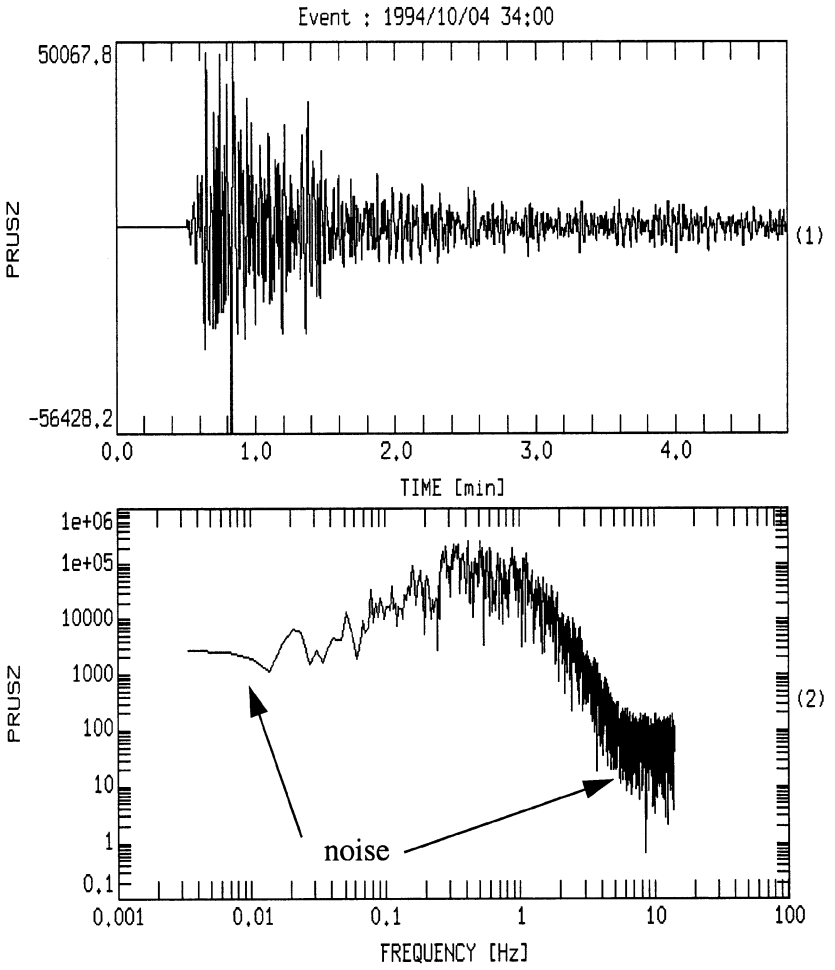



**Fig. 9.6** Vertical component record of an $m_b = 7.5$ earthquake (1994 April 4) in the Kurile islands (upper panel). The units are counts with offset removed. The lower panel shows the corresponding amplitude spectrum. Notice the noise at both the low frequency and high frequency ends of the amplitude spectrum.

In Fig. 9.7, the amplitude spectrum is compared to the displacement frequency response function of the recording system.

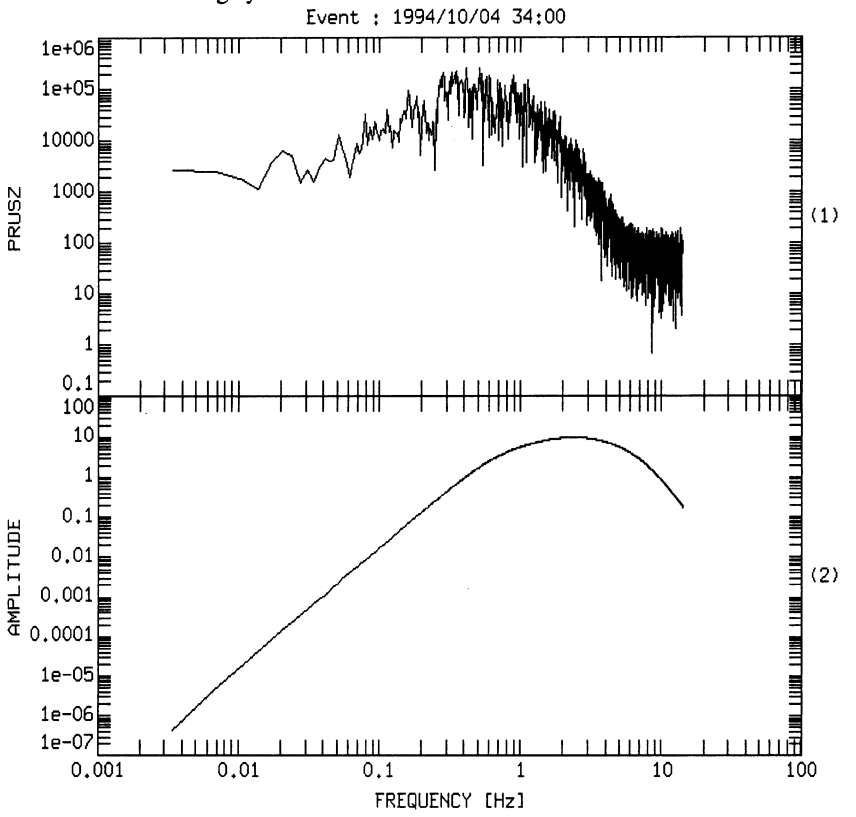



**Fig. 9.7** Amplitude spectrum from Fig. 9.6 compared to the modulus of the displacement frequency response function at station PRU (lower panel).

Notice the roll-offs of the frequency response function at low (seismometer) and high frequencies (anti-alias filter), respectively. The high frequency decay of the seismic spectrum falls within the relatively flat portion of the frequency response function and can therefore not be attributed to the instrument. From the convolution theorem we know that we can describe the interaction of a general linear time invariant filter with an arbitrary input signal by multiplying the appropriate spectra. Hence, to deconvolve the impulse response of a recording system, we can - with the proper zero padding to avoid wrap around effects - calculate the DFT spectrum of the seismic trace (top trace in Fig. 9.7) and divide it by the frequency response function (bottom trace in Fig. 9.7). This technique is called spectral division. The resulting trace and the corresponding amplitude spectrum are shown in Fig. 9.8.

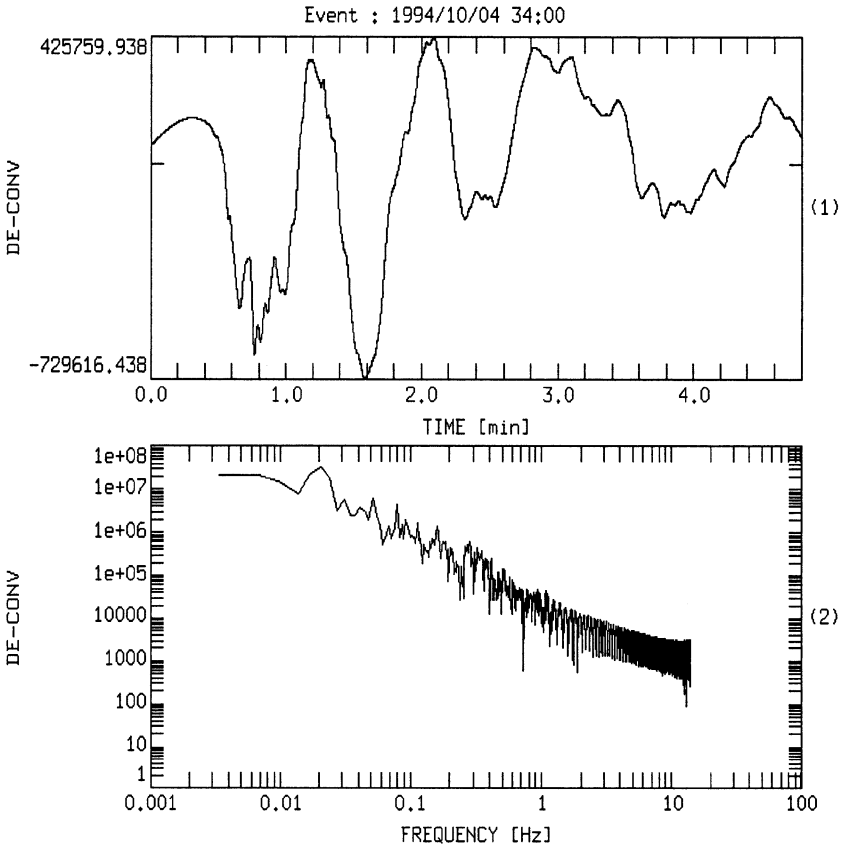



**Fig. 9.8** Deconvolution by spectral division of the spectrum shown in the top trace of Fig. 9.7 using the frequency response function displayed in the bottom trace of Fig. 9.7.

Although the deconvolution has not gone completely crazy, we can easily see that there are stability problems. The low frequency components outside the passband of the instrument dominate the amplitude spectrum. In the high frequency range we also see noise amplification similar to that sketched in Fig. 9.5. In Fig. 9.9 the deconvolution is visualized in the spectral domain. From top to bottom the figure shows the amplitude spectrum of the observed trace, the amplitude part of the frequency response of the inverse filter, and the amplitude spectrum of the deconvolved trace. The bottom trace is the complex product of the top and center trace. Fig. 9.9 again illustrates how the low and high frequency portions of the signal are disproportionately enhanced during the deconvolution process. As a consequence, the deconvolved signal may be dominated by these components.

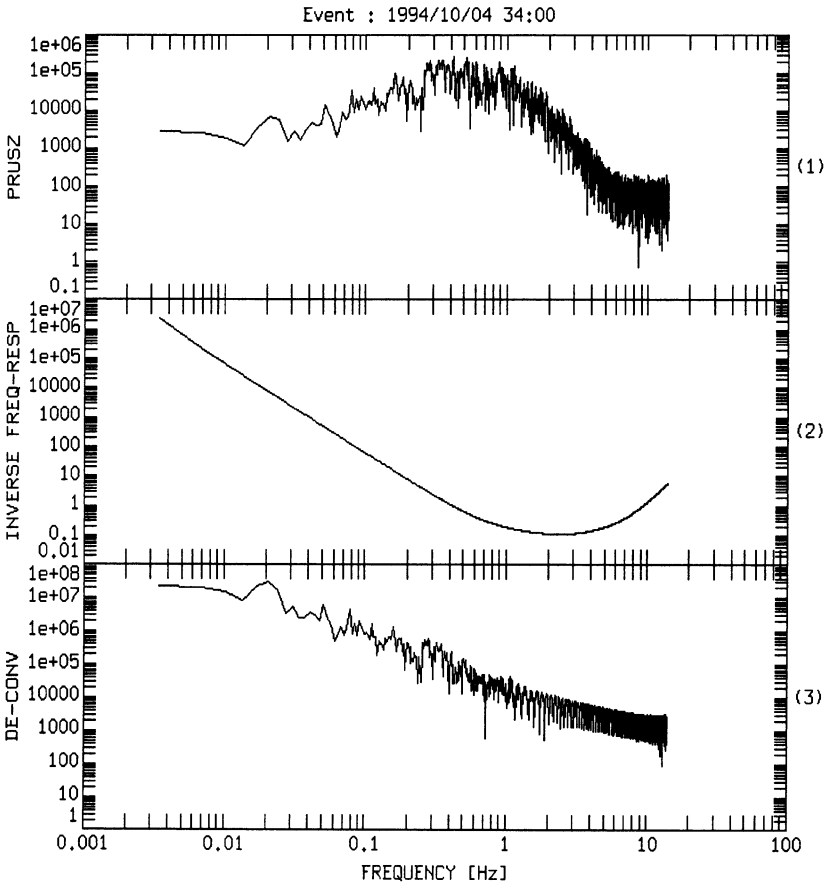



**Fig. 9.9** Deconvolution illustrated in the frequency domain. From top to bottom, the amplitude spectrum of the observed trace, the frequency response function of the inverse filter, and the amplitude spectrum of the deconvolved trace are displayed.

One way to minimize this problem is to filter the final trace with a band-pass filter, thereby restricting the signal to the valid frequency band. Another method that is commonly used with spectral division deconvolution is the so-called *water level* correction. In this method a threshold for the amplitude values of the denominator spectrum is enforced while the phase is left unchanged. The water level is measured from the maximum value of the amplitude of the denominator spectrum. In other words, we trade some of the dynamic range of the denominator spectrum for stability in the spectral division. There is no general rule for determining the best water level. The denominator spectrum should be modified only enough to ensure stability during deconvolution. In Fig. 9.10, the results of applying two different waterlevel corrections (80 and 60 dB) to the deconvolution of the signal in Fig. 9.6 are shown.

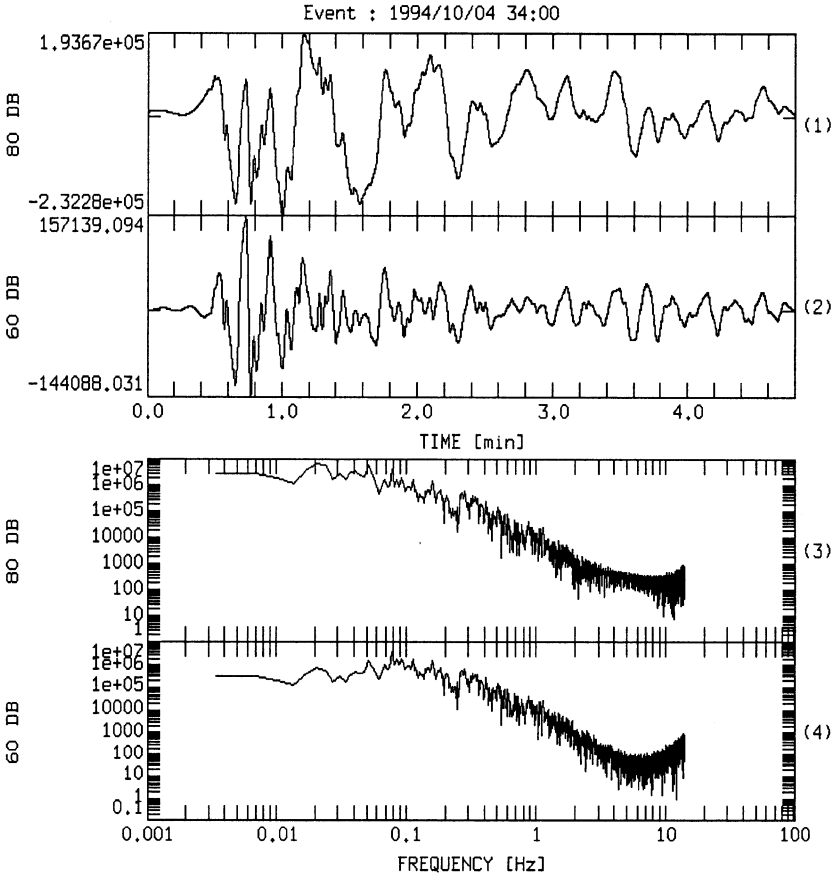



**Fig. 9.10** Deconvolution by spectral division using a waterlevel stabilization of 80 dB (top trace) and 60 dB (second trace) of the signal shown in Fig. 9.6. The corresponding amplitude spectra are shown in the two bottom traces.

Noise amplified during the deconvolution stage will be mapped into the frequency response function of the simulated instrument if the passband of the simulated instrument extends beyond the valid frequency band. This is schematically illustrated in Fig. 9.11. Fig. 9.11a shows a situation which corresponds to a case where the passband of the simulated instrument is fully within the valid frequency band. The situation in Fig. 9.11b, however, illustrates what happens if at least part of the passband of the simulated instrument lies outside the valid frequency band. In this case, the simulated signal is distorted as is sketched in Fig. 9.11b. If these effects are too strong, the simulated seismogram may be completely dominated by noise.

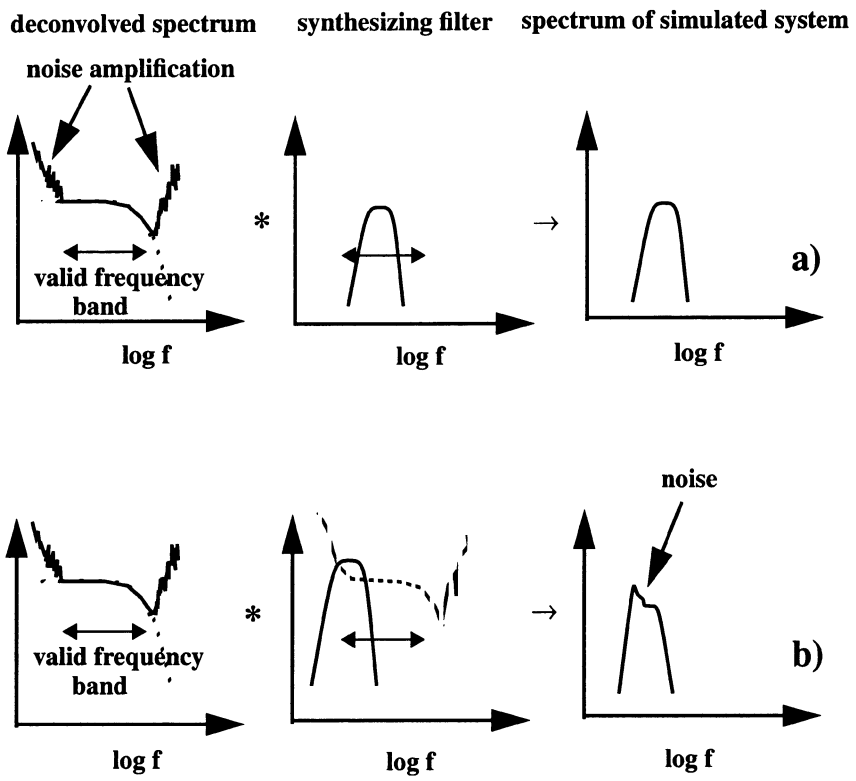



**Fig. 9.11** Mapping of deconvolution noise into the simulated system. a) The passband of the simulated instrument is fully within the valid frequency band. b) The passband of the simulated instrument is part outside the valid frequency band. Hence the deconvolution noise maps into the simulated system.

Noise is not the only source of distortion or instabilities during the simulation process. If we consider an electrodynamic seismometer system, the continuous transfer function of the deconvolution filter is of the form

$$T^{-1}(s) = \frac{s^2 + 2h\omega_0 s + \omega_0^2}{s^3} \quad . \tag{9.2}$$

Here, $\omega_o = 2\pi f_0$, with $f_0$ the undamped seismometer eigenfrequency and $h$ the damping factor. From (9.2) we see that even in the noise-free case, there is a problem for $s = 0$ (zero frequency) since $T^{-1}(s)$ becomes singular (irregular). For (9.2) *regularization* can be achieved by subsequent high pass filtering, for example with the transfer function

$$T_{HP}(s) = \frac{s^3}{(s+\varepsilon)^3} \tag{9.3}$$

The resulting transfer function will become

$$T^{-1}(s) \cdot T_{HP}(s) = \frac{s^2 + 2h\omega_0 s + \omega_0^2}{s^3} \cdot \frac{s^3}{(s+\varepsilon)^3} = \frac{s^2 + 2h\omega_0 s + \omega_0^2}{(s+\varepsilon)^3} \tag{9.4}$$

which is equivalent to replacing the triple pole at $s = 0$ by a triple pole at $s = -\varepsilon$.

*Problem 9.1* For which frequencies do you expect regularity problems for a digital recording system?

Since instabilities caused by mathematical irregularity are restricted to particular frequencies - commonly zero frequency or the Nyquist frequency- one way to improve stability is by cascading notch filters for those frequencies.

Besides noise and mathematical singularities, another issue which can be seen as part of the stability problem, is *causality*. Using equation (9.1) we can write the z-transform transfer function of the simulation filter as

$$T_{sim}(z) = \frac{T_{syn}(z)}{T_{act}(z)} \ . \tag{9.5}$$

In the most general case, $T_{act}(z)$ may correspond to a mixed phase system, hence it can be written as a product of its minimum and maximum phase parts, $T_{act}^{min}(z)$ and $T_{act}^{max}(z)$, respectively

$$T_{act}(z) = T_{act}^{min}(z) \cdot T_{act}^{max}(z) \ . \tag{9.6}$$

It then follows that

$$T_{sim}(z) = \frac{1}{T_{act}^{min}(z)} \cdot \frac{1}{T_{act}^{max}(z)} \cdot T_{syn}(z) \ . \tag{9.7}$$

At this point we need to recall some relationships between the pole and zero positions of the transfer function of a discrete system and its stability and phase properties.

- *Stability.* In order for a discrete system to have a stable causal response, all the poles must be located within the unit circle.
- *Minimum phase condition.* If, in addition, all zeros are located within the unit circle, the system is called minimum phase.
- *Maximum phase condition.* In contrast, all zeros of a maximum phase system lie outside the unit circle.

*Problem 9.2* Discuss the stability and phase properties of the terms $1/T_{act}^{min}(z)$ and $1/T_{act}^{max}(z)$ in (9.7).

There are two techniques commonly used to deal with the instability arising from the fact that $1/T_{act}^{max}(z)$ contains poles outside the unit circle. The first technique replaces all the zeros of $T_{act}^{max}(z)$ by their complex conjugate reciprocals (plus proper scaling). The modulus of the resulting transfer function $\tilde{T}_{act}^{max}(z)$ is identical to that of $T_{act}^{max}(z)$. The phase portions of $T_{act}^{max}(z)$ and $\tilde{T}_{act}^{max}(z)$ would be quite different, which means that serious phase errors are introduced into the signals for the simulated system. However, since all the zeros have been moved inside the unit circle, $\tilde{T}_{act}^{max}(z)$ is a minimum phase system and therefore has a stable inverse $1/\tilde{T}_{act}^{max}(z)$. This technique is called *unilateral stabilization.*

The second stabilization technique utilizes the fact, that although a discrete system with poles outside the unit circle does not have a stable causal impulse response, it has a stable anticausal impulse response. A stable simulation filter process is obtained if the maximum phase portion of the actual instrument is deconvolved in reverse time while the minimum phase part is deconvolved without stability problems in real time. Filtering thus becomes a two step procedure. If in the first step the maximum phase part of $T_{act}(z)$ is deconvolved in reverse time, which is equivalent to replacing $z$ by $1/z$, we obtain

$$Y_{temp}(1/z) = \frac{Y_{act}(1/z)}{T_{act}^{max}(1/z)} \,. \tag{9.8}$$

Replacing $z$ by $1/z$ causes all singularities outside the unit circle to move inside. There-

fore, $1/T_{act}^{max}(1/z)$ has all its poles and zeros inside the unit circle and filtering with $1/T_{act}^{max}(1/z)$ is stable in reverse time. The second step then combines the deconvolution of the minimum phase part of $T_{act}(z)$ and the synthesizing filter $T_{syn}(z)$. Both processes are stable in regular time. The simulated signal becomes

$$Y_{sim}(z) = \frac{T_{syn}(z)}{T_{act}^{min}(z)} \cdot Y_{temp}(z) \quad . \tag{9.9}$$

Using this approach, the impulse response function of the simulation filter has anti-causal and causal components, i.e., it becomes two-sided. This technique, which can only be used on recorded signals of finite duration, is called *bilateral stabilization.*

## 9.2 Implementation

While the discussion of the stability problem can be kept fairly general, the methods used to implement the filtering process depend primarily on the information that is available about the transfer function of the recording system. This issue is closely related to the data format, since different data formats provide different options for representing instrument data (see e. g. GSE Wave Form Data Format, 1990; SEED reference manual, 1990). Since the GSE format, which we have used several times already, offers a sensible compromise in terms of power and complexity, the following discussion is restricted to those representations of calibration data which the GSE format provides. These are

- Poles and zeros of the Laplace transform transfer function (PAZ)
- FIR filter coefficients (FIR)
- Frequency-amplitude-phase triplets (FAP)
- Combinations of PAZ and FIR (FPZ)

The implementation of stable and accurate instrument simulation filters under general conditions is a very complex task. There is no single best solution which performs well under all conceivable conditions (see also *Proceedings of the seminar on deconvolution of seismograms and high-fidelity seismometry*, J. Geophys, 39, 501-626, 1973). To date, the most advanced software package for automatically performing instrument simulation and deconvolution is the program PREPROC written by Miroslav Zmeskal (Zmeskal and Plesinger, 1995; Plesinger et al., 1995). It requires the calibration information be given in GSE format and provides several methods for implementing the filter process. In this

chapter we will describe some of the theoretical background of the techniques used in PREPROC. In some respect, we are entering more advanced topics of filter implementation. The emphasis is not to enable you to write computer codes such as PREPROC but to better judge the failure or success of a particular simulation filtering process.

### 9.2.1 FROM THE SPECTRAL TO THE TIME DOMAIN

From the convolution theorem we know that the implementation of the deconvolution and/or simulation filter can be done in either the time or the frequency domain. The spectral division was an example of a frequency domain implementation. If data traces are very long, however, frequency domain techniques may require the calculation of very long DFT sequences, unless special blocking schemes are used. This is not only very time consuming; it may simply become impossible if memory is limited by the computer's operating system (e.g. DOS). For this reason, implementation of the deconvolution/ simulation filter in the time domain is often preferred.

Since the discrete transfer functions $T(z)$ with which we are dealing will always be rational functions in $z$, it is convenient to perform the actual filtering by using the difference equation. Let us briefly recall this relationship by starting from a general discrete LTI system described by its difference equation:

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{l=0}^{M} b_l x[n-l] \quad . \tag{9.10}$$

The discrete transfer function $T(z)$ was defined as the z-transform of the output $\mathcal{Z}\{y[n]\}$ divided by the z-transform of the input $\mathcal{Z}\{x[n]\}$. Taking the z-transform of (9.10) and using the shifting theorem( $x[n-n_0] \Leftrightarrow z^{-n_0}X(z)$ ), we obtain

$$\sum_{k=0}^{N} a_k z^{-k} Y(z) = \sum_{l=0}^{M} b_l z^{-l} X(z) \tag{9.11}$$

and

$$T(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{l=0}^{M} b_l z^{-l}}{\sum_{k=0}^{N} a_k z^{-k}} = \left(\frac{b_0}{a_0}\right) \frac{\prod_{l=1}^{M} (1 - c_l z^{-1})}{\prod_{k=1}^{N} (1 - d_k z^{-1})} \tag{9.12}$$

where $c_l$ are the non-trivial zeros and $d_k$ are the non-trivial poles of $T(z)$. Recalling the correspondence between a time shift of $k$ samples in time and the multiplication with $z^{-k}$ of the corresponding z-transform (shifting theorem), we can see that the coefficients of the polynomials in the transfer function $T(z)$ are identical to the coefficients of the difference equations. Hence, once we know the coefficients of the numerator and denominator polynomials of the transfer function we can construct the difference equation for filtering (9.10).

In the special case of (9.11) for $N = 0$ and $a_0 = 1$, the denominator polynomial vanishes and the transfer function is completely represented by its zeros $c_k$.

$$\begin{aligned} T(z) = \frac{Y(z)}{X(z)} &= \sum_{l=0}^{M} b_l z^{-l} = b_0 \prod_{l=1}^{M} (1 - c_l z^{-1}) \\ &= z^{-M} \cdot b_0 \prod_{l=1}^{M} (z - c_l) \end{aligned} \tag{9.13}$$

Since the transfer function in (9.12) has only trivial poles (at $z = 0$), this filter is always stable and the difference equation reduces to

$$y[n] = \sum_{l=0}^{M} b_l x[n-l] \quad . \tag{9.14}$$

This corresponds to the convolution of the finite filter coefficient sequence $b[l]$ for $l = 0, ...., M$ and the time sequence $x[n]$. For $x[n] = \delta[n]$ the response of the filter is equal to the sequence $b[l]$. Hence, $b[l]$ is called a finite impulse response function or a FIR filter.

Two additional points are worth mentioning in this context. First, the stability of the FIR filter process does not necessarily guarantee the stability of the corresponding inverse of a FIR filter, and secondly, the statement that FIR filters do not contain poles, is not necessarily true for the origin or infinity. We have already discussed this property - which can be seen from the shifting theorem ( $x[n-n_0] \Leftrightarrow z^{-n_0} X(z)$ ) - in the context of Problem 7.3.

### 9.2.2 FIR SIMULATION FILTERS

The implementation of the simulation filter as a FIR filter is an approach that is very general and works for all types of instrument representations within the GSE format. The filter coefficients are calculated by inverse DFT of the frequency response function of the simulation filter. The calculation of the FIR simulation filter coefficients can be conceptually separated into:

- *The calculation of the frequency response functions of the recording instrument and the instrument to be synthesized.* If the DFT is used, the frequency response function must be evaluated at equally spaced points in the frequency domain. For the FAP format, this is commonly done by interpolation between frequency-amplitude-phase triplets. If the frequency response function is given in terms of the poles and zeros of the continuous transfer function $T(s)$ (PAZ format), the frequency response function at equally spaced frequencies can be obtained directly by evaluating $T(s)$ at $s = k \cdot j2\pi\Delta f$ with $-N/2 < k \leq N/2$ and $\Delta f = f_{dig}/N$. Here $N$ is the number of samples taken for the DFT. If the FFT algorithm is used, $N$ must be a power of 2. If an instrument is represented in terms of its corresponding FIR filter coefficients (FIR format), the corresponding frequency response function is obtained by DFT.
- *The calculation of the frequency response function of the simulation filter.* In this step, the frequency response function $T_{sim}(j\omega) = T_{syn}(j\omega)/T_{act}(j\omega)$ is evaluated by complex division. This step may include any of the stabilization measures we have discussed.
- *The calculation of the FIR coefficients by inverse DFT.* After inverse DFT we obtain $N$ filter coefficients. We should keep in mind that the DFT for $N$ points corresponds to the Fourier series representation of an infinite periodic sequence which repeats after $N$ samples. $N$ must be chosen large enough to ensure a good frequency resolution and that the inverse DFT at sample $N$ has decayed to negligible amplitudes. Otherwise we encounter the problem of wrap around effects, a temporal equivalent of the alias effect.

*The choice of the filter length*

The wrap around effect (which we have already encountered in a different context chapter 7.8) is schematically illustrated in Fig. 9.12. An easy way to avoid this effect is to choose $N$ sufficiently large. On the other hand, for filtering using the difference equation (9.10), a large number of filter coefficients is very inefficient and not desirable. The tricky part of any FIR filter implementation is to decide on the actual length of the filter to be used and *how* to truncate the $N$ coefficients obtained from the inverse DFT to a reasonable number of $M$ coefficients to be used during filtering (9.10).

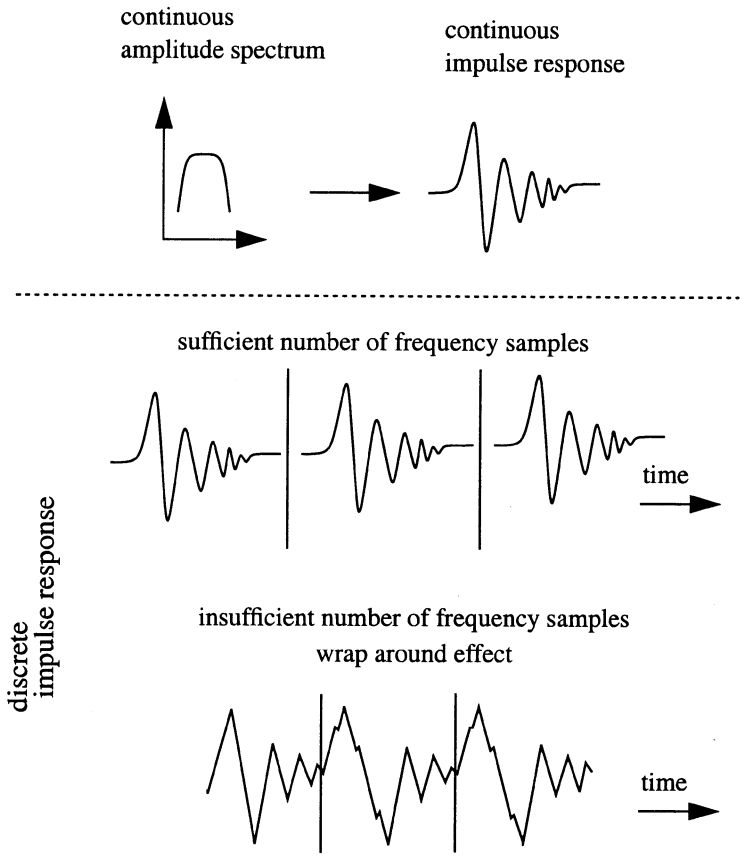



**Fig. 9.12** Temporal wrap around effect during FIR filter implementation when an insufficient number of points is used for the inverse DFT.

The effect of FIR filter length - which is not restricted to instrument simulation filters- is demonstrated below by a simple example of a pure synthesizing filter. In Fig. 9.13, the spectrum of a FIR filter approximating the velocity frequency response function of station A1 of the GRF array is shown for filters of three different lengths. From top to bottom the lengths are N = 32, 64, and 128. In each case the data were zero padded to 1024 points.

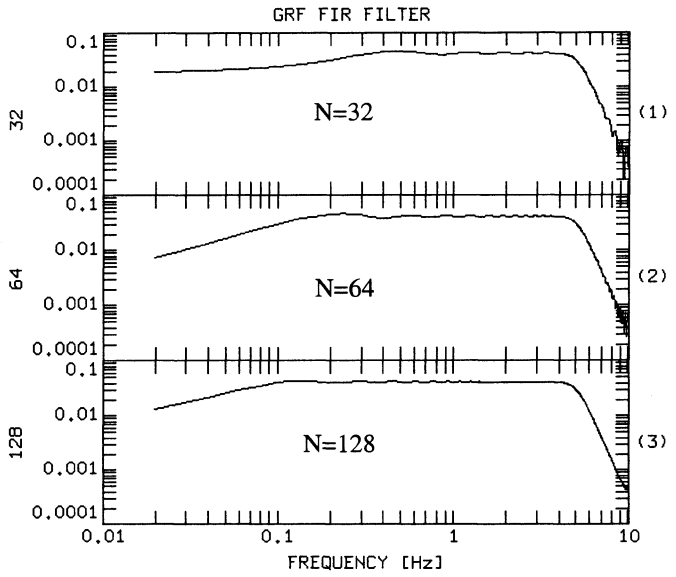



**Fig. 9.13** DFT spectra for the velocity frequency response function for different filter lengths.

The original impulse response function was calculated for a spike at the first sample of a trace of 1024 points and a sampling rate of 20 Hz. Note the spurious ripples in the frequency response function, especially noticeable in the passband between 0.05 and 5 Hz. They are obviously related to the number of FIR filter coefficients used for the calculation since the effect decreases when the number of coefficients increases. The oscillatory behaviour of the frequency response functions of the truncated filters can be understood from the frequency domain properties of a boxcar window as is illustrated in Fig. 9.14.

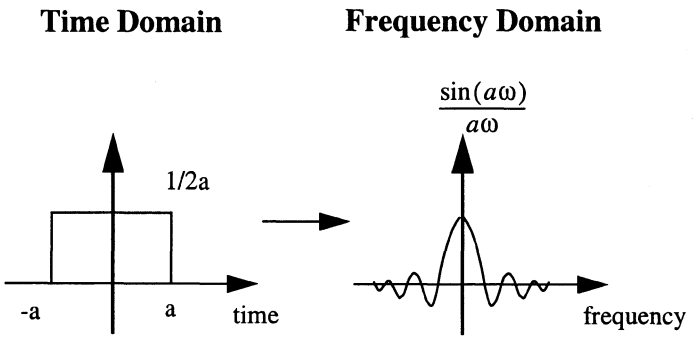



**Fig. 9.14** Time and frequency domain properties of a boxcar tapering window.

Truncation in the time domain corresponds to multiplication with a boxcar window. From the convolution theorem we know that the spectrum of the truncated signal equals the spectrum of the original signal convolved with the spectrum of the boxcar window ('multiplication in the time domain' corresponds to 'convolution in the frequency domain'). The spectrum of the boxcar window, however, is a $\sin(a\omega)/(a\omega)$ function if $2a$ is the length of the boxcar window. This is a highly oscillatory function which causes the ripples in Fig. 9.13. One way to decrease the amplitude of the oscillations is to taper the margins of the truncation window. This corresponds to smoothing the spectrum.

### 9.2.3 BILINEAR TRANSFORM

If the transfer function of the recording system is given in terms of the poles and zeros of the analog system (PAZ format), it is convenient to approximate the corresponding discrete transfer function $T(z)$ from $T(s)$ by a mapping procedure called the bilinear transform. This is defined as (e.g., Oppenheim and Willsky, 1989)

$$s = \frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}} \tag{9.15}$$

with $T$ the sampling interval and $s$ and $z$ the variables of the complex s- and z-planes, respectively. The inverse transform is given by

$$z = \frac{1 + (T/2)s}{1 - (T/2)s} \tag{9.16}$$

The values of the discrete transfer function $T(z)$ become

$$T(z) = T(s)\Big|_{s = (2/T)[(1 - z^{-1})/(1 + z^{-1})]} \quad . \tag{9.17}$$

In chapter 7.6, we introduced the $z$-transform as the discrete equivalent of the Laplace transform, in other words as the Laplace transform of a continuous function sampled at intervals $T$ (see chapter 7.6). The corresponding mapping between the s-plane and the z-plane was described by $z = e^{sT}$ and $s = ln(z)/T$ and was shown in Fig. 7.3. The mapping of the bilinear transform is different as is demonstrated in Fig. 9.15.

The s-plane representation in eqn. (9.15) can be interpreted as taking the first term of the expansion of the natural logarithm into continuous fractions:

$$s = \frac{1}{T} \cdot \ln(z) = \frac{1}{T} \cdot \frac{2(z-1)}{z+1-\ldots} \approx \frac{2}{T} \cdot \frac{z-1}{z+1} = \frac{2}{T} \cdot \frac{1 - z^{-1}}{1 + z^{-1}} \text{ (see equation (9.15)).}$$

The same mapping equation is also obtained for the discrete time simulation of the integration process using the trapezoidal rule (Kraniauskas, 1992).

If we set $s = \sigma + j\omega$ in equation (9.16) we obtain:

$$z = \frac{[2/T + \sigma] + j\omega}{[2/T - \sigma] - j\omega} \tag{9.18}$$

If $\sigma = 0$, $|z| = 1$, hence the imaginary axis is mapped onto the unit circle. For $\sigma > 0$ the modulus of the numerator is always greater than the modulus of the denominator, therefore $|z| > 1$. Thus, the right half s-plane is mapped onto the outside of the unit circle in the z-plane. Finally, from the same argument we can see that the left half-plane of the complex s-plane is mapped onto the interior of the unit circle in the z-plane. Since this is true for all the poles and zeros, stable continuous systems will be mapped onto stable discrete systems, and the phase properties of the continuous systems will remain the same in the corresponding discrete system.

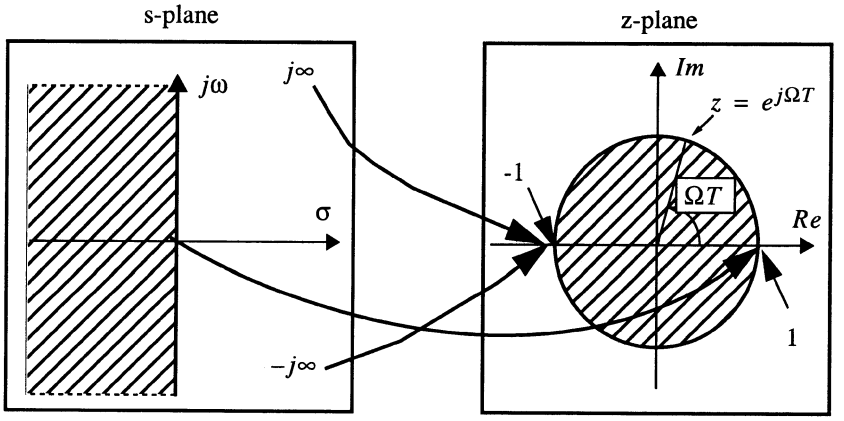



**Fig. 9.15** Mapping of the complex s-plane onto the complex z-plane using the bilinear transform.

From (9.15) we obtain for $z = e^{j\Omega T}$

$$s = \frac{2}{T} \cdot \frac{1 - e^{-j\Omega T}}{1 + e^{-j\Omega T}} = \frac{2}{T} \cdot \frac{2e^{j\frac{\Omega T}{2}}(j\sin(\Omega T/2))}{2e^{j\frac{\Omega T}{2}}\cos(\Omega T/2)} = j\frac{2}{T} \cdot \tan(\Omega T/2) = j\omega. \tag{9.19}$$

Thus, the discrete frequency $\Omega$ $(-\pi/T \le \Omega \le \pi/T)$ maps onto the continuous frequency $\omega$ $(-\infty$ to $\infty)$ according to

$$\omega = \frac{2}{T} \cdot \tan(\Omega T/2) \tag{9.20}$$

and correspondingly

$$\Omega = \frac{2}{T} \operatorname{atan}(\omega T/2) . \tag{9.21}$$

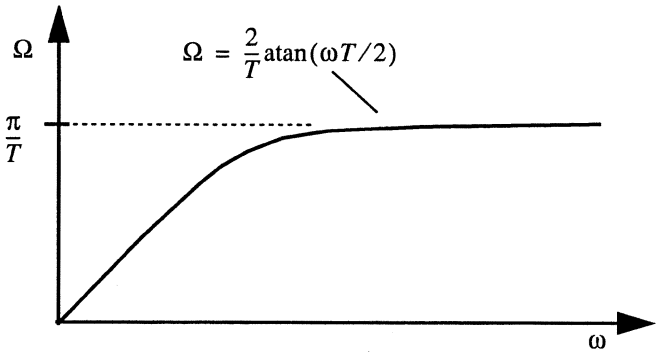



**Fig. 9.16** Mapping of the continuous frequency $\omega$ onto the discrete frequency $\Omega$.

Fig. 9.16 shows this relationship graphically. Zero continuous frequency $\omega = 0$ is mapped onto $\Omega = 0$ and hence $s = 0$ maps onto $z = 1$ (see also Fig. 9.15). Infinite frequency $\omega = \infty$ maps onto $\Omega = \pi/T$ and therefore $s = j\infty$ is mapped onto $z = e^{j\pi} = -1$. For $s = -j\infty$ we obtain $\Omega = -\pi/T$ and again $z = e^{-j\pi} = -1$.

The bilinear transform maps the complete continuous frequency axis onto the unit circle in the z-plane. However, the price for this convenience is a non-linear distortion of the frequency axis. For example a lowpass filter with a corner frequency of $\omega_1$ in the continuous transfer function will have a corner frequency of $\Omega_1 = \frac{2}{T} \operatorname{atan}(\omega_1 T/2)$ in the discrete transfer function which is lower than desired. This is demonstrated in Fig. 9.17 for a 3 pole Butterworth lowpass filter with a corner frequency of 1 Hz and a sampling frequency of 5 Hz.

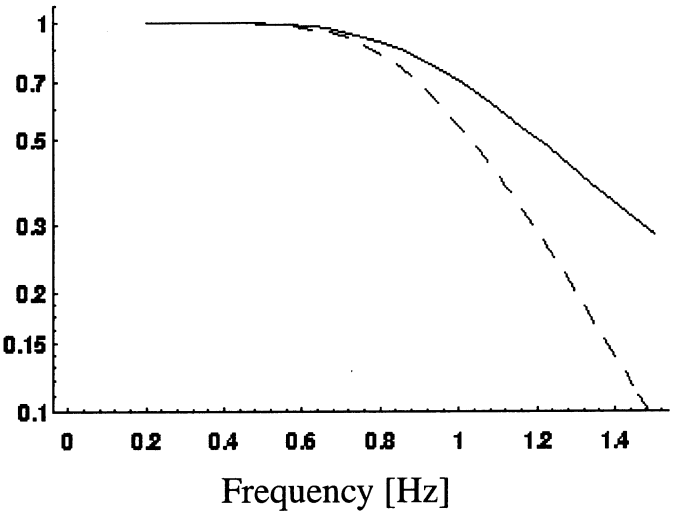



**Fig. 9.17** Modulus of the continuous and discrete frequency response function for a three pole Butterworth filter with an analog corner frequency of 1Hz. The continuous line shows the analog frequency response while the dashed line shows the discrete frequency response. The discrete corner frequency (modulus equals $1/\sqrt{2}$) is at 0.893 Hz.

To compensate for this distortion, we start with a *warped* corner frequency $\omega'_1 = \frac{2}{T} \cdot \tan(\omega_1 T/2)$. According to (9.21) this maps to

$$\Omega = \frac{2}{T}\text{atan}\left(\frac{2}{T} \cdot \tan(\omega_1 T/2)T/2\right) = \omega_1 \quad , \tag{9.22}$$

the desired value. Thus, if the continuous transfer function consists of regions of approximately constant slopes which change at certain critical frequencies, we can compensate for the frequency distortion during the bilinear transformation by warping all the critical frequencies $\omega_c$ into $\omega'_c$ using the relationship

$$\omega'_c = \frac{2}{T} \cdot \tan(\omega_c T/2) \tag{9.23}$$

and then performing the bilinear transform using these warped frequencies. It should be noted, however, that in order for this strategy to be successful, all the critical frequencies must be less than the Nyquist frequency $1/(2T)$. For the 1 Hz lowpass filter shown in

Fig. 9.17 the warped corner frequency for a sampling frequency of 5 Hz is 1.1563 Hz. The resulting frequency response function is displayed in Fig. 9.18.

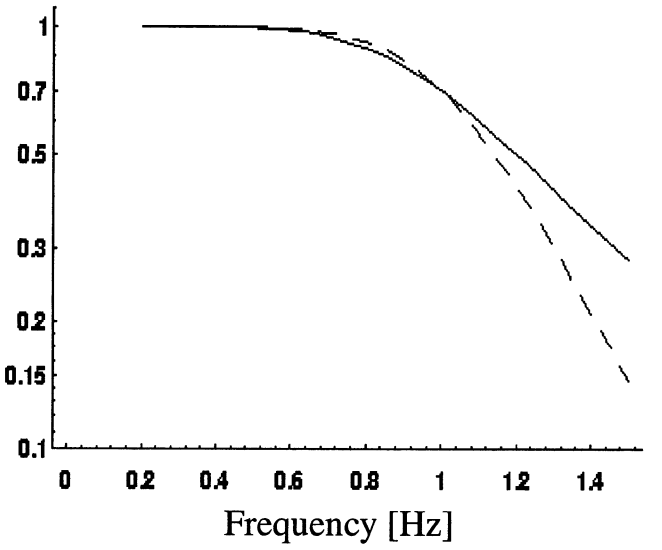



**Fig. 9.18** The effect of corner frequency warping on the filter displayed in Fig. 9.17. The continuous line shows the analog frequency response while the dashed line shows the discrete frequency response with corner frequency warping. Both analog and discrete corner frequency equal 1 Hz.

Let us assume a seismogram has been recorded by an analog instrument which can be described using the transfer function $T(s) = s^2/(s^2 + 2h\omega_{01}s + \omega_{01}^2)$ with $h_1 = 0.707$ and $\omega_{01} = 2\pi \cdot 0.1[Hz]$ , and then been digitized with a sampling frequency of 20 Hz. We wish to create a simulation filter which changes the seismometer eigenfrequency to $\omega_{02} = 2\pi \cdot 0.5[Hz]$ and the damping constant to $h_2 = 0.5$ using the bilinear transform. For this task we can ignore the properties of the anti-alias filter, so the continuous transfer function for the simulation filter becomes

$$T_{sim}(s) = \frac{(s^2 + 2h_1\omega_{01}s + \omega_{01}^2)}{(s^2 + 2h_2\omega_{02}s + \omega_{02}^2)} . \tag{9.24}$$

Replacing $\omega_{01}$ and $\omega_{02}$ by their warped equivalents $\omega'_{01}$ and $\omega'_{02}$ using (9.23) we obtain

$$T_{sim}(s) = \frac{(s^2 + 2h_1\omega'_{01}s + \omega'^2_{01})}{(s^2 + 2h_2\omega'_{02}s + \omega'^2_{02})} \tag{9.25}$$

The discrete transfer function of the simulation filter is calculated by replacing $s$ in (9.25) by $2 \cdot (1 - z^{-1})/(T \cdot (1 + z^{-1}))$ and by replacing $\omega'_{01}$ and $\omega'_{02}$ with $\frac{2}{T} \cdot \tan(\omega_{01}T/2)$ and $\frac{2}{T} \cdot \tan(\omega_{02}T/2)$, respectively. The result is

$$T_{sim}(z) = \frac{\left(\frac{1-z^{-1}}{1+z^{-1}}\right)^2 + 2h_1\tan(\omega_{01}T/2)\left(\frac{1-z^{-1}}{1+z^{-1}}\right) + \tan^2(\omega_{01}T/2)}{\left(\frac{1-z^{-1}}{1+z^{-1}}\right)^2 + 2h_2\tan(\omega_{02}T/2)\left(\frac{1-z^{-1}}{1+z^{-1}}\right) + \tan^2(\omega_{02}T/2)} \tag{9.26}$$

Expanding this expression into negative powers of $z$ gives

$$T_{sim}(z) = \frac{b_0 + b_1 \cdot z^{-1} + b_2 \cdot z^{-2}}{a_0 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2}} \tag{9.27}$$

with

$$\begin{aligned}
a_0 &= 1 + 2h_2\tan(\omega_{02}T/2) + \tan^2(\omega_{02}T/2) \\
a_1 &= 2\tan^2(\omega_{02}T/2) - 2 \\
a_2 &= 1 - 2h_2\tan(\omega_{02}T/2) + \tan^2(\omega_{02}T/2) \\
b_0 &= 1 + 2h_1\tan(\omega_{01}T/2) + \tan^2(\omega_{01}T/2) \\
b_1 &= 2\tan^2(\omega_{01}T/2) - 2 \\
b_2 &= 1 - 2h_1\tan(\omega_{01}T/2) + \tan^2(\omega_{01}T/2)
\end{aligned} \tag{9.28}$$

For the values of $h_1, \omega'_{01}, h_2$ and $\omega'_{02}$ given above, the coefficients become

$a_0 = 1.0849, a_1 = -1.98761, a_2 = 0.927492$ and

$b_0 = 1.02246, b_1 = -1.99951, b_2 = 0.978034$.

Using equations (9.10) through (9.12) and dividing by $a_0$, the filter difference equation is

$$\begin{aligned} y[n] &= 1.83208 \cdot y[n-1] - 0.854914 \cdot y[n-2] \\ &\quad + 0.94245 \cdot x[n] - 1.84304 \cdot x[n-1] \\ &\quad + 0.9015 \cdot x[n-2] \end{aligned} \tag{9.29}$$

Note the sign change of the coefficients $a_1$ and $a_2$ in (9.29). The performance of this filter in the time and frequency domains is demonstrated with synthetic data in Fig. 9.19.

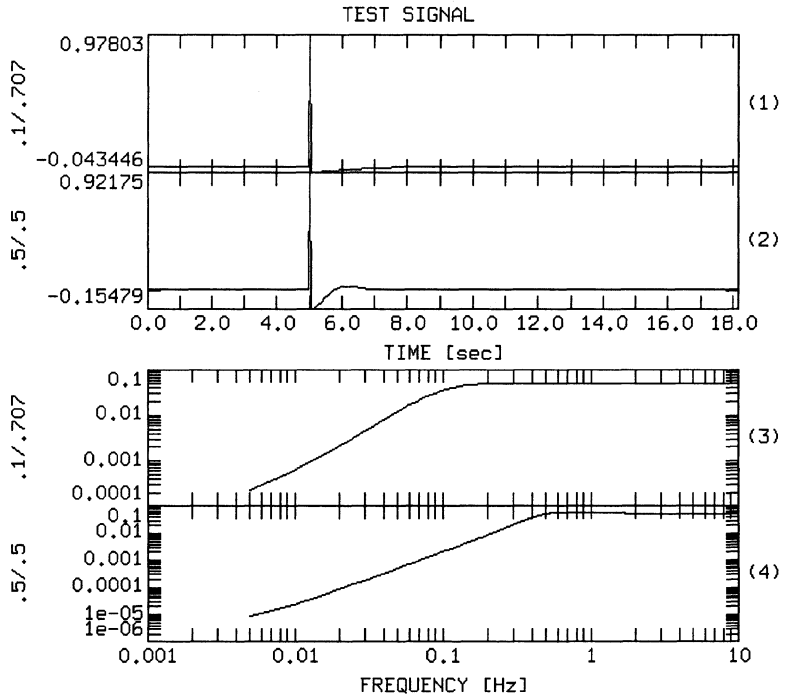



**Fig. 9.19** Changing the eigenfrequency and damping factor of digital records by recursive simulation filtering using the bilinear transform. Panels 1, 3 and panels 2, 4 show the impulse response function and the modulus of the frequency response function before and after the simulation filtering, respectively.

The bilinear transform is a powerful method for implementing an instrument simulation filter as a recursive filter if the continuous transfer function is known as a rational function (e.g., if the poles and zeros are given). Using this approach we can simulate general combinations of seismometer-galvanometers from broadband seismograms as demonstrated by Seidl (1980). Due to the nonlinear frequency distortion, however, it can only

be done if the frequency response function of the system to be simulated can be divided into frequency bands in which the slope of the response function is approximately constant and the cutoff frequencies between these bands are much lower than the Nyquist frequency.

*Problem 9.3* Use the bilinear transformation with frequency warping to design a recursive filter to simulate a 'displacement' seismometer with a seismometer eigenfrequency of 1/120 Hz, a damping factor of 0.707, and a sampling frequency of 20 Hz.

*Problem 9.4* Determine the difference equation for a deconvolution filter which completely removes the instrument defined in Problem 9.3.

*Problem 9.5* Determine the difference equation for a deconvolution filter for an electrodynamic sensor with the same eigenfrequency, damping and sampling frequency as in Problem 9.3.

### 9.2.4 LEAST SQUARES APPROACH

Another very general method for implementing time-domain recursive simulation filters is to approximate their complex frequency response function using the complex frequency response function of a recursive filter. The desired frequency response function for the simulation filter $T_{des}(j\omega_n)$ can be calculated using the same approach as for the FIR filter, from any of the GSE formats. For a recursive filter with $M$ non-trivial zeros $\hat{c}_l$ and $N$ non-trivial poles $\hat{d}_k$, the frequency response function $T_{recur}(j\omega_n)$ is obtained by evaluating

$$T_{recur}(z) = \frac{\sum_{l=0}^{M} \hat{b}_l z^{-l}}{\sum_{k=0}^{N} \hat{a}_k z^{-k}} = \left(\frac{\hat{b}_0}{\hat{a}_0}\right) \frac{\prod_{l=1}^{M} (1 - \hat{c}_l z^{-1})}{\prod_{k=1}^{N} (1 - \hat{d}_k z^{-1})} \tag{9.30}$$

on the unit circle ($z = e^{j \cdot \omega_n T}$) for discrete frequencies $\omega_n = n \cdot 2\pi/(NDFT \cdot T)$ and $n = 0...NDFT$. Here, $T$ is the sampling interval in seconds. The approximation error between $T_{des}(j\omega_n)$ and $T_{rec}(j\omega_n)$ is

$$e(\omega_n) = T_{des}(j\omega_n) - T_{recur}(j\omega_n) \; . \tag{9.31}$$

The principle of the least squares approach is to find a transfer function $T_{recur}(z)$ which minimizes the total approximation error $E = \sum_n |e(\omega_n)|^2$. This requires the solution of a non-linear optimization problem for which a number of different strategies exist. The discussion of these, however, is beyond the scope of this text. They are discussed for example in Parks and Burrus (1987). It should be mentioned that the filter design problem can also be formulated as a least squares problem in the time domain. In this case the approximation error has to be calculated from the difference between the actual and the desired filter impulse response function in the time domain. In addition to the actual optimization problem, a major difficulty in applying the least squares approach is the order selection, i.e., selecting the numbers of poles and zeros in (9.30). For orders which are too low, the approximation will be very poor, while orders which are too high result in spurious oscillations of $T_{recur}(z)$.

Like the bilinear transformation approach, the least squares methods results in recursive simulation filters which operate in the time domain. In contrast to the bilinear transform, it performs well for high cut-off frequencies, that is, for cases where the valid frequency band comes close to the Nyquist frequency. On the other hand, for highly oversampled data it is inferior to the bilinear transformation method.

## 9.3 Review inverse and simulation filtering

In this chapter we have discussed methods and strategies for obtaining stable and accurate instrument simulation filters. The variety of potential problems illustrates that this is a complex task. Nevertheless, Miroslav Zmeskal (Zmeskal and Plesinger, 1995) have implemented all the procedures in this chapter within a single software package (PREPROC) which performs instrument simulation and deconvolution more or less automatically. As a matter of fact, one of the intentions of this chapter was to provide the theoretical background to PREPROC within the framework of the concepts discussed in this text. A prerequisite for the application of PREPROC is that the instrument information is provided in the form of GSE calibration files.

One of the most important lessons of this chapter is that there is no single solution to the instrument simulation problem which performs well in all conceivable situations (see also *Proceedings of the seminar on deconvolution of seismograms and high-fidelity seismometry*, J. Geophys, 39, 501-626, 1973). The bilinear transform, for example, performs better than the least squares approach for highly oversampled data while for non-oversampled data the opposite is true. Besides the performance issue, the method of choice depends also on the kind of information available on both the actual recording system and the system to be simulated. While the bilinear transformation approach is restricted to the simulation of continuous transfer functions given in poles and zeros, the two most general methods in that respect are the FIR filter method and the least squares method

which can both be used with all the formats provided by the GSE convention.

Since the simulation can become potentially unstable, we must be aware of various techniques which can be employed for stabilization. There are different causes of instability, among them noise and mathematical irregularity. Consequently there are also different cures. Besides restricting the frequency band for the simulation to the valid frequency band, mathematical irregularities can be treated by post-filtering with appropriate notch filters. Furthermore, causality can be traded for stability in the context of bilateral stabilization.

Finally it should be noted that the deconvolution problem can also be viewed as a linear inverse problem and can be treated within the framework of inverse theory. Since this topic is beyond the scope of this text, interested readers are referred to the papers by Oldenburg (1981), Treitel and Lines (1982), O'Dowd (1990), Sipkin and Lerner-Lam (1992), and Gurrola et al. (1995).

CHAPTER 10

# The measurement of wavelet parameters from digital seismograms

Standard observational measurements from seismograms only include a relatively small number of signal parameters such as onset times, amplitudes, rise times, pulse widths, signal moments and onset polarities of particular wave groups (Fig. 10.1).

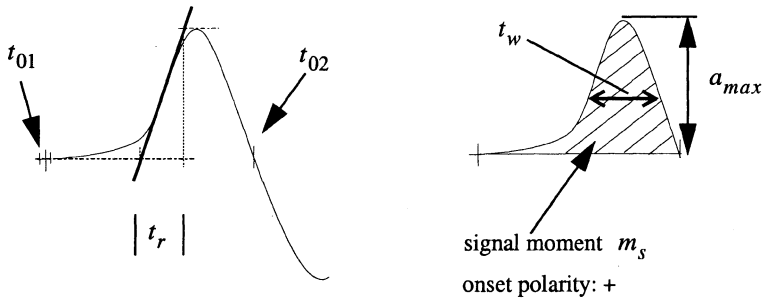



**Fig. 10.1** Schematic waveform onset and commonly determined signal parameters such as onset time, $t_{01}$, onset polarity, amplitude, $a_{max}$, rise time, $t_r$, pulse width, $t_w$ or time of zero crossing, $t_{02}$, and signal moment $m_s$.

Onset times, for example, are essential for solving the earthquake location problem and are used extensively for structural studies such as velocity tomography. Measurements of amplitude and signal moment are most relevant for the determination of magnitudes and seismic moments, respectively. Pulse durations and rise times contain information about source parameters such as source size and stress drop, as well as, about the attenuation structure of the earth, while onset polarities are commonly used to determine focal mechanisms. Most important in the present context is the fact that each of these parameters can be severely affected by the properties of seismic recording systems (Fig. 10.2).

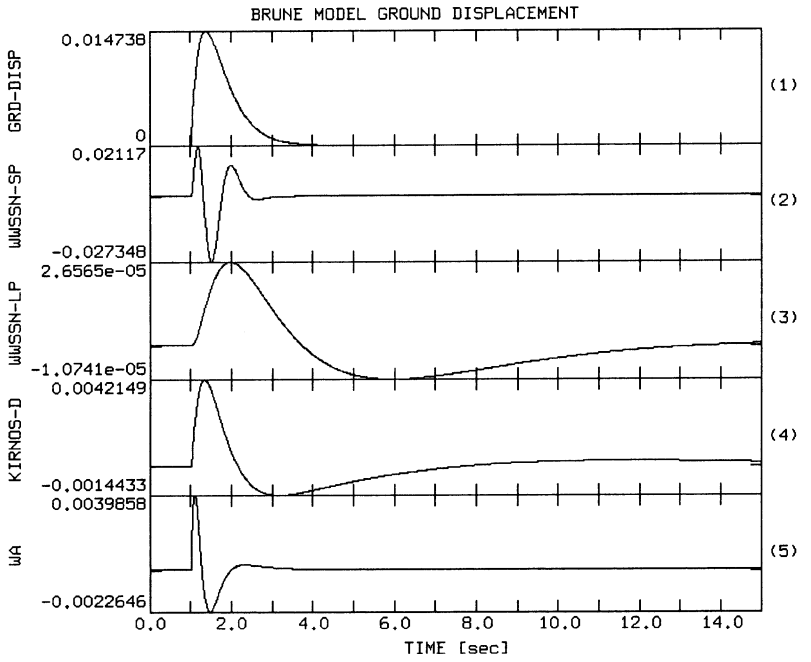



**Fig. 10.2** Distortion of a synthetic ground displacement signal (top trace) by standard seismograph systems (traces 2-5). Displayed in traces 2 to trace 5 are the output signals of simulated WWSSN-SP, WWSSN-LP, KIRNOS, and a Wood-Anderson instrument for the synthetic input signal shown in trace 1. For a collection seismogram examples see Kulhanek (1990).

Even a casual inspection of Fig. 10.2 demonstrates that the waveform parameters, except for the onset polarity, determined directly from a seismogram may yield values which have little resemblance to the parameters which represent the true ground motion. The situation becomes even worse if we imagine a realistic scenario in which additional noise is present in the data traces.

In the following we will discuss in greater detail the determination of individual waveform parameters and the assumptions commonly made in this context. One reason for this discussion is to increase the awareness of the influence of the properties of recording systems on seismologically relevant parameters. A consequence will be a more quantitative understanding of the limitations instrumental properties can potentially impose on the interpretation of waveforms and hence the geophysical models which we derive from them.

## 10.1 The determination of ground motion amplitudes

Although one of the most important tasks of standard observatory data processing, the determination of ground motion (displacement, velocity, or acceleration) amplitudes from digital seismograms is often a matter for confusion. Body and surface wave magnitudes, for example, both require the estimation of the amplitude of ground displacement (Willmore, 1979). Although in principle not very complicated, the transformation from digital counts into ground motion amplitudes contains various pitfalls and deserves very careful treatment. Part of the confusion is caused by sloppy use of the terms transfer function and frequency response function. Since the term transfer function is defined as the ratio of two Laplace transforms, we must be very precise in defining a transfer function of a certain type, such as a *displacement transfer function* or a *velocity transfer function*, in order to avoid one of the pitfalls. Another source of confusion is the assignment of ground motion units to digital seismograms. Strictly speaking, the unit of a digital seismogram is counts! We can only talk about a well - defined correspondence between a ground motion amplitude and a digital amplitude in counts under certain restrictions, e.g. if we are dealing with a harmonic signal in steady state. We often speak of a 'velocity system' whenever the sensor operates on the basis of a moving coil, forgetting that the proportionality of the output to ground velocity depends on a number of factors such as the frequency bandwidth of the signal, the damping properties of the seismometer/galvanometer combination, and the frequency passband of the recording system. This is not merely a theoretical consideration: moving coil seismometers with heavily overdamped galvanometers (e.g. the Kirnos instrument; Willmore, 1979), produce seismograms which are proportional to ground displacement within a certain frequency band. We will see that, especially for recording systems with frequency passbands much narrower than the bandwidth of the input signal of interest, we could save ourselves many headaches if we would use counts as the only unit for the digitized output signal.

*Problem 10.1* Fig. 10.3 shows three signals with different signal character and frequency content. The corresponding spectra are shown in Fig. 10.4. Discuss qualitatively what happens to these signals if they are recorded with the instrument for which the amplitude frequency response function is shown in Fig. 10.5. Under which circumstances would it be justified to call this a 'velocity recording system'?

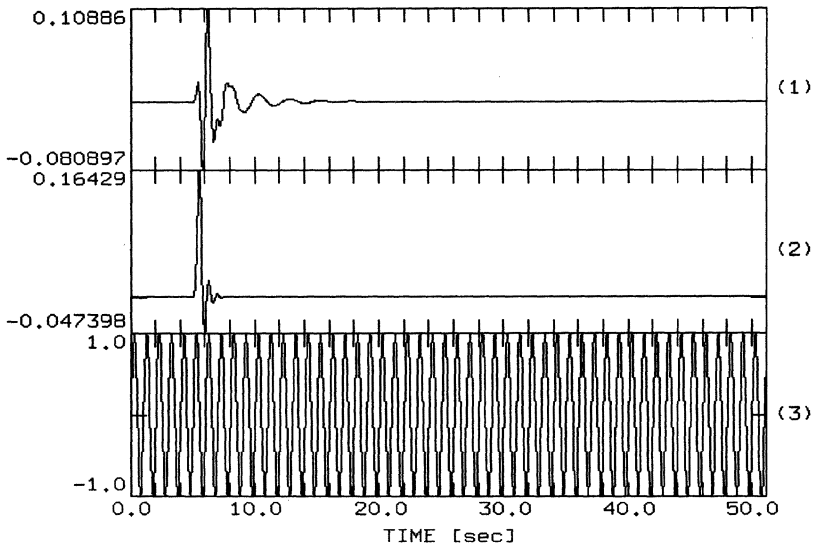



**Fig. 10.3** Example input signals to the seismograph system whose amplitude frequency response is shown in Fig. 10.5.

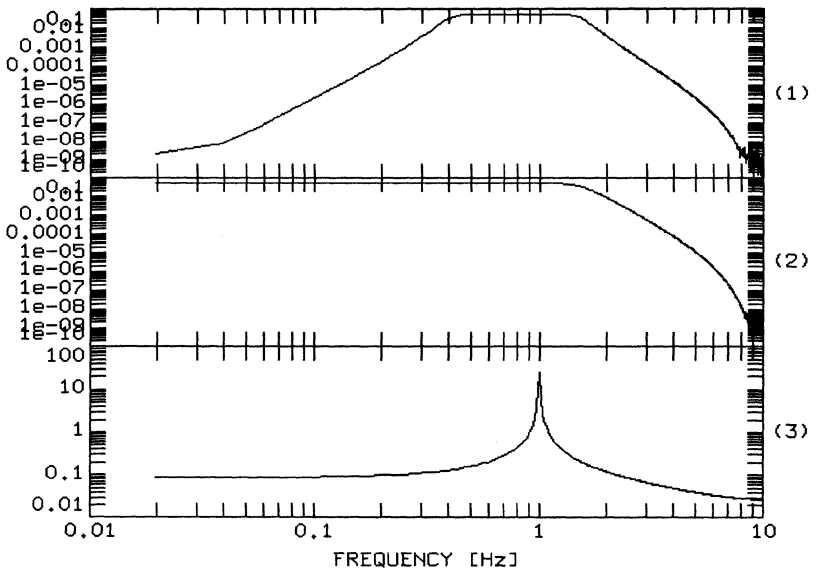



**Fig. 10.4** DFT spectra for the signals in Fig. 10.3.Note that the spectrum of the sinusoid shows the effects of the boxcar tapering window used.

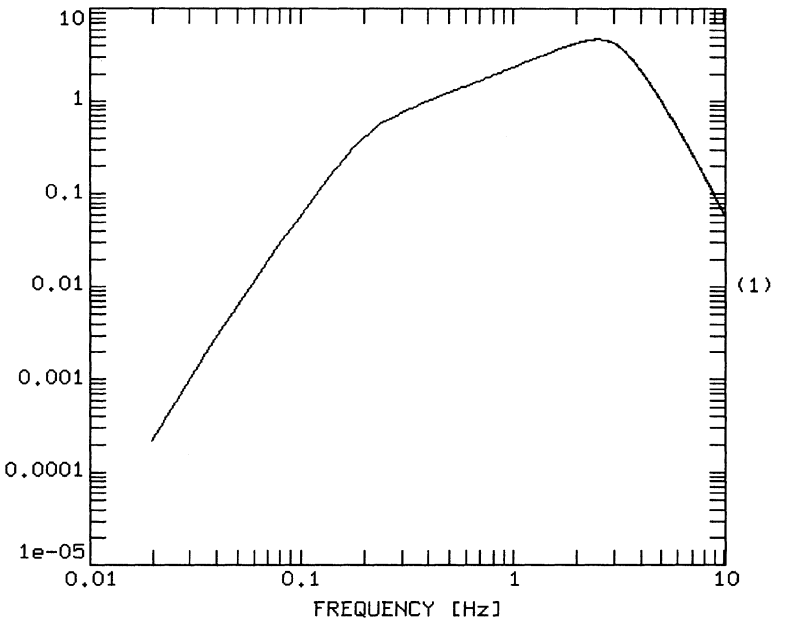



**Fig. 10.5** Amplitude frequency response function for a fictitious recording system (displacement magnification curve).

Amplitude frequency response functions such as in Fig. 10.5 are often referred to as ***magnification curves*** for the corresponding instrument (e.g. Willmore, 1979). Let us back up a little bit to see how this fits into the concepts we have encountered so far.

The frequency response function of a general filter has been introduced as the Fourier transform of the output signal divided by the Fourier transform of the input signal.

$$T(j\omega) = \frac{\mathcal{F}\{Output\}}{\mathcal{F}\{Input\}} = \frac{Output(j\omega)}{Input(j\omega)} \tag{10.1}$$

The filter process can be described very conveniently in the frequency domain. The Fourier transform of the output signal is obtained by multiplying the Fourier transform of the input signal by the frequency response function. For an electrodynamic sensor, the output signal would be a voltage, while for a digital system, the output signal would be a sequence of digits. Depending on whether the **input signal** is in ground displacement, ground velocity, or ground acceleration, we are speaking of a *displacement-, velocity-*, or *acceleration frequency response function*, respectively.

The following rule is a fairly easy way to remember the units properly:

- Given the output signal of a seismic recording system (in whatever physical units it provides), deconvolution by the displacement impulse response will yield ground displacement, deconvolution by the velocity impulse response will yield ground velocity, and deconvolution by the acceleration impulse response will yield ground acceleration.

*Problem 10.2* What are the proper units for a displacement-, a velocity-, and an acceleration frequency response function of a digital recording system (including seismometer)?

For an impulsive input signal (Dirac impulse), the output signal of the filter was called the impulse response function. Since in this case the Fourier transform of the input signal is unity, we can see from equation (10.1) that the frequency response function is the Fourier transform of the impulse response function. In this context, we also use the terms displacement-, velocity-, and acceleration impulse response functions depending on whether the input signal (impulse) is a ground displacement, ground velocity, or ground acceleration impulse.

An obvious and fairly safe way to reduce confusion in the use of the terms displacement-, velocity-, and acceleration response is to use only a single type. This strategy is followed in the context of the GSE format which uses only the displacement response. If we know one type of response function we can obtain the others by using simple operations. These are given by the relationships between integration and differentiation in the time domain and the corresponding operations in the frequency domain which we have already used. The following problem discusses such a calculation.

*Problem 10.3* Given the velocity impulse response function for a seismic recording system, how can one calculate the corresponding displacement- and acceleration impulse response functions in both time- and frequency domain?

An important property of linear systems useful in this context is the fact that the values of the frequency response functions are eigenvalues of the system with the corresponding eigenvectors given by the harmonic functions. In a linear transformation represented by a linear equation $\vec{x} \rightarrow A \cdot \vec{x}$, an eigenvector is transformed onto itself $\vec{x} \rightarrow \lambda \cdot \vec{x}$; hence its direction is not changed. By analogy harmonic signals as 'eigenvectors' do not change their frequency content (Fig. 10.6).

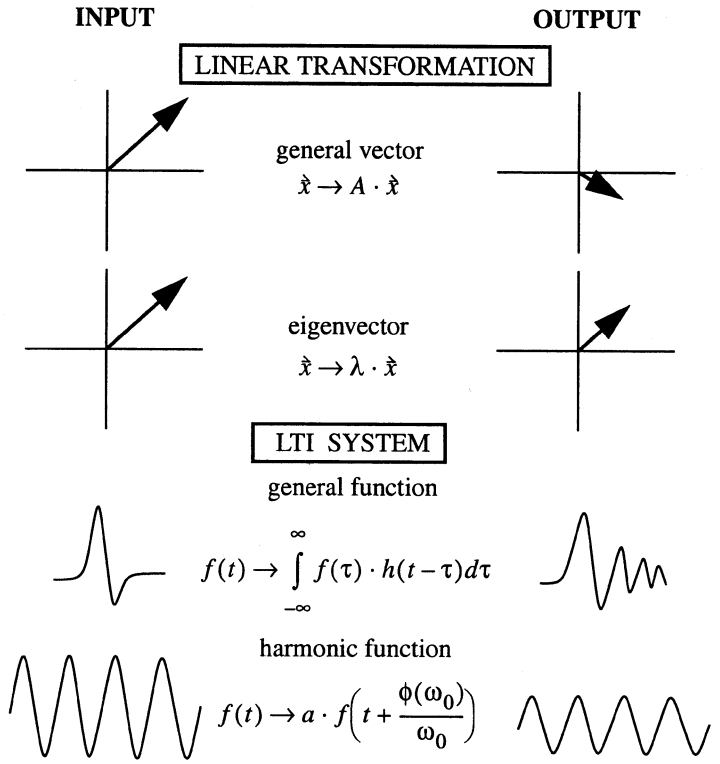



**Fig. 10.6** Eigenvector concept in linear transformations and LTI systems.

Making use of this concept, we can easily understand what we have already seen in Problem 10.1 for trace 3. An harmonic input signal:

$$A_o \sin(\omega_0 t) \tag{10.2}$$

filtered by a system described by the frequency response function

$$T(j\omega) = |T(j\omega)| \cdot e^{j\Phi(\omega)} \tag{10.3}$$

does not change its frequency, only its amplitude and phase. The filter output signal is

$$|T(j\omega_0)| A_o \sin(\omega_0 t + \Phi(\omega_0)) \ . \tag{10.4}$$

Hence, the modulus of the frequency response function is exactly the amount of *magnification* which a harmonic input signal will experience. This explains the use of the name magnification curves for plots such as Fig. 10.5. For short period instruments, a technique commonly used to measure the frequency response function is to put the sensor on a shaking table which is then excited by harmonic signals of different frequencies. From the *magnification* and phase shift for each frequency, the complete frequency response function can be approximated. For broadband instruments, the transfer function is identified by different means (e.g. Plesinger, 1993).

For the determination of body and surface wave magnitudes, it is common to convert digital counts into ground motion amplitudes by determining both signal amplitude and period and subsequently dividing the signal amplitude in counts by the value of the amplitude of the displacement frequency response function (magnification curve) at the signal frequency. PITSA (since version 4.0) uses this method which reproduces the way analog records are commonly analysed. Although this is the correct procedure for monochromatic harmonic signals, for the type of broadband signals normally encountered in earthquake seismology, it is accompanied by very strong assumptions.

In earlier chapters we saw how the frequency response function $T(j\omega)$ can be obtained from the more general concept of the transfer function $T(s)$ by letting $s = j\omega$, which corresponds to evaluating $T(s)$ on the imaginary axis of the complex plane. The transfer function of a general continuous Nth order LTI system can be written in the form

$$T(s) = \frac{\beta_0 + \beta_1 s + \beta_2 s^2 + \ldots + \beta_L s^L}{\alpha_0 + \alpha_1 s + \alpha_2 s^2 + \ldots + \alpha_N s^N} = \left(\frac{\beta_L}{\alpha_N}\right) \frac{\prod_{k=1}^{L} (s - s_{0k})}{\prod_{k=1}^{N} (s - s_{pk})} \tag{10.5}$$

with $s_{0k}$ and $s_{pk}$ the k-th zero and pole of the transfer function, respectively. For a given distribution of poles and zeros we can therefore express the frequency response function $T(j\omega)$ as

$$T(j\omega) = \left(\frac{\beta_L}{\alpha_N}\right) \frac{\prod_{k=1}^{L} (j\omega - s_{0k})}{\prod_{k=1}^{N} (j\omega - s_{pk})} \tag{10.6}$$

A similar expression is valid for discrete systems if we use the transfer function concept defined on the basis of the z transform. In any case, the impulse response function is

obtained by the corresponding inverse transform.

For a classical inertial seismometer with a displacement transducer the transfer function $T(s)$ is of the form (cf. equation (4.52))

$$T_{disp}(s) = -c\frac{s^2}{s^2 + 2h\omega_0 s + \omega_0^2} \tag{10.7}$$

with $\omega_o = 2\pi f_0$ ($f_0$ = undamped seismometer eigenfrequency), $h$ the damping factor, and $c$ a frequency - independent scale factor. **The scale factor $c$ has the units of the output signal divided by the unit of the input signal while the frequency - dependent portion is assumed to be dimensionless.** The two poles of the transfer function are the roots of the denominator polynomial

$$s_{p(1,2)} = -(h \pm \sqrt{h^2 - 1}) \cdot \omega_0 \ . \tag{10.8}$$

Since the numerator polynomial is simply $s^2$ there are only two zeros at the origin of the complex plane. Using the correspondence between differentiation in the time domain and multiplication by $s$ in the $s$ domain, we arrive at the displacement transfer function for an electrodynamic seismometer (inertial seismometer with a velocity transducer) by simply multiplying equation (10.7) by $s$.

$$T_{disp}(s) = -c'\frac{s^3}{s^2 + 2h\omega_0 s + \omega_0^2} \ . \tag{10.9}$$

If the output signal is in $V$, the scale factor $c'$ in this case has the unit $[V/m]$.

In Problem 10.3 we saw that the displacement frequency response function can be obtained from the velocity frequency response function by multiplication with $j\omega$, or in terms of the transfer functions by multiplication with $s$. Therefore, in order to obtain the corresponding velocity transfer function for the system described in (10.7), we have to divide (10.7) by $s$. Hence **the displacement transfer function for a classical inertial seismometer given in (10.7) is of the same mathematical form as the velocity transfer function for an electrodynamic system.**

$$T_{vel}(s) = -G\frac{s^2}{s^2 + 2h\omega_0 s + \omega_0^2} \tag{10.10}$$

The modulus of the corresponding velocity frequency response function (magnification curve) of an electrodynamic system is schematically plotted in Fig. 10.7.

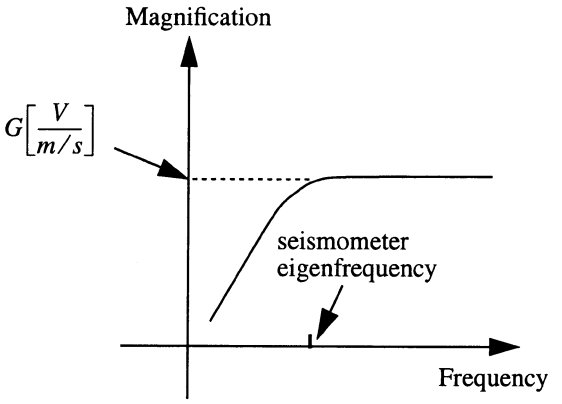



**Fig. 10.7** Sketch of the velocity magnification curve of an electrodynamic system.

For frequencies above the seismometer eigenfrequency the magnification for ground velocity is constant. The output voltage of an electrodynamic seismometer for frequencies above the eigenfrequency is directly proportional to ground velocity. The proportionality constant (scale factor) in this case is commonly called the generator constant $G$ and is given in $\left[\frac{V}{m/s}\right]$. In other words, the generator constant describes the magnification - or modulus of the frequency response function - for a ground velocity input signal for frequencies above the seismometer eigenfrequency.

*Problem 10.4* Determine the displacement frequency response function for an electrodynamic seismometer system with eigenfrequency $f_0 = 0.0082706$ Hz, $h = 0.718$ and a generator constant of $G = 1500$ [V/m/s]. Use equation (10.7) and convert the resulting transfer function into the required form. In addition, express the frequency response function by the roots of the transfer function.

While equation (10.7) allows us to relate the output voltage of an electrodynamic seismometer to the input ground velocity, when using digital data we need to relate counts to ground motion. Hence, we must know how the output voltage of the seismometer was amplified and sampled by the ADC. In other words, we need to know the voltage corresponding to a single count, a quantity which is called the LSB (least significant bit) of the ADC (cf. chapter 6.1).

*Problem 10.5* Calculate the displacement frequency response function relating ground motion [nm] to [counts] for the seismometer in Problem 10.4 given an LSB of 2.5 $\mu V$.

*Problem 10.6* In order to practice the conversion of frequency response functions one more time, convert the displacement frequency response function from Problem 10.5 into the corresponding velocity response function. Discuss the various possible approaches.

In many cases - although not always- it is possible to express the transfer function of the overall recording system using poles and zeros and an appropriate scale factor. In this case, the frequency response function can be expressed as

$$T(j\omega) = C \cdot \frac{\prod_{k=1}^{L} (j\omega - s_{0k})}{\prod_{k=1}^{N} (j\omega - s_{pk})} = C \cdot F(j\omega) \; . \tag{10.11}$$

If the frequency response function in (10.11) is a displacement frequency response function, the scale factor $C$ is exactly that given in the calibration section of the GSE format to assure proper scaling of amplitudes. Depending on the information available about the instrument, $C$ can be calculated in various ways. In Problem 10.4 we obtained it for the velocity frequency response of an electrodynamic system using the generator constant.

In a more general scenario, let us assume we only know the poles and zeros of the system and the magnification of a harmonic signal in steady state at a single certain frequency. In terms of GSE conventions this is referred to as the *calibration gain* in [ground motion per digital count] at a given calibration period. It corresponds to an amplitude which could be measured from a shaking table experiment. Although the frequency dependence of the calibration gain is not expressed explicitly, it is important to keep it in mind. Let us start out with a situation where the input signal is given in ground displacement while the output signal is given in counts ('displacement response') as sketched in Fig. 10.8.

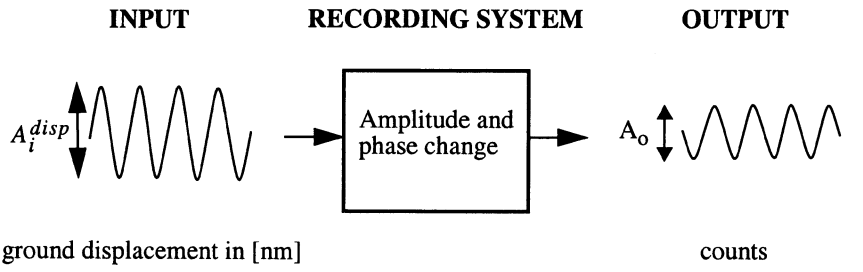



**Fig. 10.8** Sketch of a shake table experiment to measure the calibration gain $g_d$ at a given period.

If, for the situation sketched in Fig. 10.8, $A_i^{disp}$ is the peak to peak amplitude of the input signal in nm and $A_o$ the corresponding output amplitude in counts, the calibration gain $g_d$ at the calibration frequency would be $g_d = A_i^{disp}/A_o$ in [nm/counts]. The relationship between the magnification curve for displacement (amplitude portion of the frequency response function), calibration gain, and calibration frequency is sketched in Fig. 10.9.

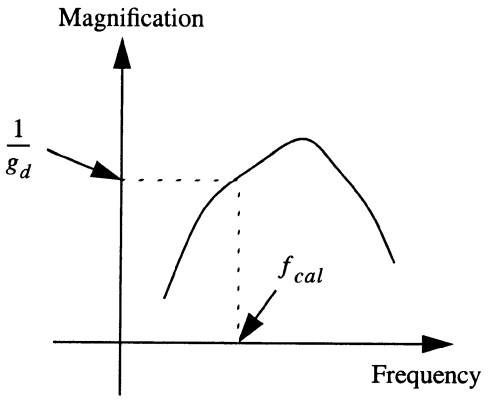



**Fig. 10.9** Calibration of a displacement amplitude frequency response function (displacement magnification curve).

The absolute value or modulus of the displacement frequency response function at the calibration frequency $f_{cal}$ is $1/g_d$. This naturally leads to the following problem.

*Problem 10.7* For a displacement frequency response function $T_{disp}(j\omega) = C_{disp} \cdot F_{disp}(j\omega)$, calculate the scale factor $C_{disp}$ in (10.11) from the knowledge of $g_d$ and the calibration period $T_{cal}$.

Although the transfer function and frequency response function in the GSE convention always refers to displacement input signals, for practical reasons, GSE convention also allows that the calibration gain be given in units of motion other than ground displacement, namely ground velocity or acceleration. In these cases, we call the corresponding gain factors $g_v$ for velocity and $g_a$ for the acceleration calibration gain. A current in the calibration coil of a seismometer, for example would correspond to ground acceleration. Within the framework of a hypothetical shake table experiment, these situations correspond to a harmonic shake table motion measured not in ground displacement, but in ground velocity for $g_v$, or in ground acceleration for $g_a$. We could simply imagine measuring the shake table excitation with a velocity- or acceleration sensor in these cases. Since we have often made use of the correspondence between differentiation in the time domain and multiplication with $s$ or $j\omega$ in the complex domain, we should be able to solve the following problem.

*Problem 10.8* Given the displacement calibration gain, $g_d$, at the calibration period $T_{cal}$, derive the relationships between $g_a$, $g_v$, and $g_d$, and give the scale factor $C_{disp}$ in (10.11) in terms of one of these quantities.

In addition to the scale factor $C$ in (10.11) and the calibration gains $g_a$, $g_v$, and $g_d$, there are two other related terms that are sometimes used in discussions of amplitude measurements from digital data. These are the *channel sensitivity* $S$ and the *normalization factor* $N$.

The channel sensitivities, $S_d$, $S_v$, and $S_a$, for displacement, velocity, and acceleration, respectively, describe the sensitivity of the system (in counts per ground motion unit) at some normalization frequency $\omega_{norm}$ within its passband. The channel sensitivity is simply the modulus of the frequency response function (or magnification curve) at the normalization frequency. Formally we can write

$$S_d = |T_{disp}(j\omega_{norm})| = C_{disp} \cdot |F_{disp}(j\omega_{norm})| \tag{10.12}$$

$$S_v = |T_{vel}(j\omega_{norm})| = \frac{|T_{disp}(j\omega_{norm})|}{\omega_{norm}} = C_{disp} \cdot \frac{|F_{disp}(j\omega_{norm})|}{\omega_{norm}} . \tag{10.13}$$

and

$$S_a = |T_{acc}(j\omega_{norm})| = \frac{|T_{disp}(j\omega_{norm})|}{\omega^2_{norm}} = C_{disp} \cdot \frac{|F_{disp}(j\omega_{norm})|}{\omega^2_{norm}} . \quad (10.14)$$

Solving for $C_{disp}$ we obtain from (10.12) - (10.14)

$$\begin{aligned} C_{disp} &= S_d \cdot \frac{1}{|F_{disp}(j\omega_{norm})|} \\ &= S_v \cdot \frac{\omega_{norm}}{|F_{disp}(j\omega_{norm})|} \\ &= S_a \cdot \frac{\omega^2_{norm}}{|F_{disp}(j\omega_{norm})|} \end{aligned} \quad (10.15)$$

The normalization factors, $N_d$, $N_v$, and $N_a$, for displacement, velocity, and acceleration, are the reciprocals of the modulus of the frequency - dependent term $F(j\omega)$ in (10.11) at the normalization frequency, $\omega_{norm}$, for displacement-, velocity-, and acceleration frequency response functions, respectively

$$N_d = \frac{1}{|F_{disp}(j\omega)|} \quad (10.16)$$

$$\begin{aligned} N_v &= \frac{1}{|F_{vel}(j\omega_{norm})|} = \frac{\omega_{norm}}{|F_{disp}(j\omega_{norm})|} \\ &= \omega_{norm} \cdot N_d \end{aligned} \quad (10.17)$$

$$\begin{aligned} N_a &= \frac{1}{|F_{acc}(j\omega_{norm})|} = \frac{\omega_{norm}}{|F_{vel}(j\omega_{norm})|} = \frac{\omega^2_{norm}}{|F_{disp}(j\omega_{norm})|} \\ &= \omega_{norm} \cdot N_v = \omega^2_{norm} \cdot N_d \end{aligned} \quad (10.18)$$

In GSE conventions we only have to deal with the frequency response function given for displacement $T(j\omega)^{GSE} \equiv T_{disp}(j\omega) = C_{disp} \cdot F_{disp}(j\omega)$ . Therefore, below we focus on the determination of $C_{disp}$ in terms of the quantities we know. If we insert (10.16) -

(10.18) into (10.15) we can express $C_{disp}$ in terms of channel sensitivities and normalization factor as

$$C_{disp} = S_d \cdot N_d = S_v \cdot N_v = S_a \cdot N_a. \tag{10.19}$$

If we let $\omega_{norm} = \omega_{cal}$, the channel sensitivities $S_d$, $S_v$, and $S_a$ equal the reciprocals of the corresponding calibration gains $g_d$, $g_v$, and $g_a$, respectively, and we can express the displacement frequency response function as

$$\begin{aligned} T_{disp}(j\omega) &= C_{disp} \cdot F_{disp}(j\omega) \\ &= S_d \cdot N_d \cdot F_{disp}(j\omega) = \frac{1}{g_d} \cdot N_d \cdot F_{disp}(j\omega) \\ &= S_v \cdot N_v \cdot F_{disp}(j\omega) = \frac{1}{g_v} \cdot N_v \cdot F_{disp}(j\omega) \\ &= S_a \cdot N_a \cdot F_{disp}(j\omega) = \frac{1}{g_a} \cdot N_a \cdot F_{disp}(j\omega) \end{aligned} \tag{10.20}$$

If the frequency - dependent term $F_{disp}(j\omega)$ of the displacement frequency response is multiplied by the displacement normalization factor $N_d$, its modulus is unity at the normalization frequency $\omega_{norm}$. Equation (10.20) describes how the displacement frequency response function can be calculated from the frequency-dependent term $F_{disp}(j\omega)$ and any reasonable combination of calibration gain, normalization factor and channel sensitivity. The frequency-dependent term can be obtained from the poles and zeros of the transfer function by using (10.11).

Expressions similar to (10.20) for $T_{vel}(j\omega)$ and $T_{acc}(j\omega)$ are easily derived using the correspondences mentioned earlier. Since we are primarily concerned with the GSE convention, however, we restrict ourselves to considering the displacement frequency response function.

The following problem gives an example of how one can derive the transfer function for an instrument if the seismometer eigenfrequency and damping factor, the generator constant, the amplification factor, and the LSB are known.

*Problem 10.9* Imagine a sensor for which the output voltage is proportional to ground velocity for frequencies above 1 Hz. The sensor output (generator constant 100 V/m/s) is amplified by a factor of 250 before being fed into a 20 bit ADC with an LSB of 1 $\mu V$. Calculate the values of $g_v$ and $g_d$ for calibration frequencies of 5 and 10 Hz, respectively. Calculate the velocity- and displacement frequency response functions for a damping factor of 0.7. Where do $g_v$ and $g_d$ plot in these figures?

Now we should be able to actually determine ground motion amplitudes from digital seismograms. An example is given in Problem 10.10.

*Problem 10.10* Imagine the following signal has been recorded on the instrument defined in Problem 10.9. What would be corresponding peak-to-peak amplitude of the ground displacement in nm?

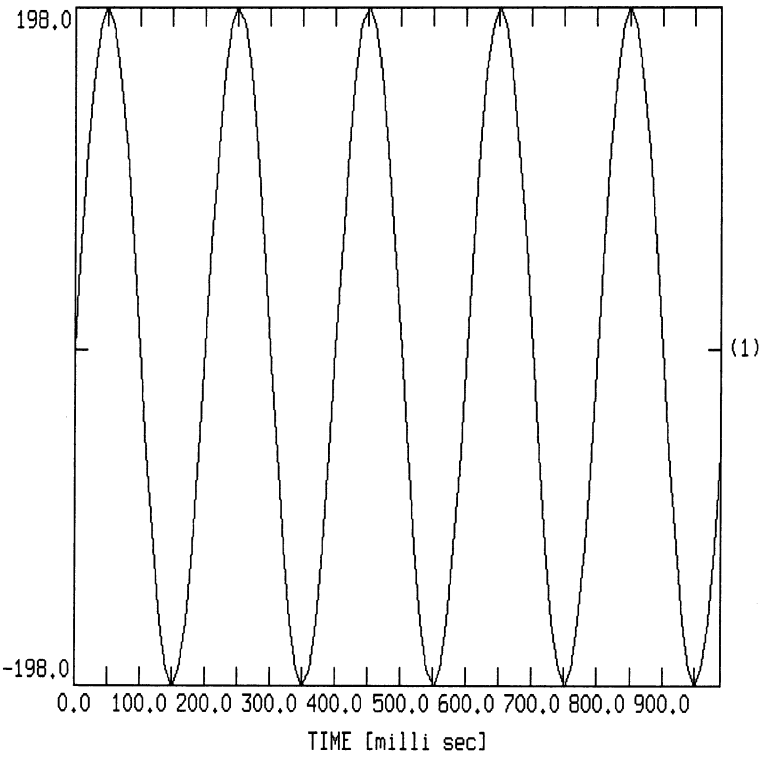



**Fig. 10.10** Signal recorded with the instrument defined in Problem 10.9.

### 10.1.1 REVIEW AMPLITUDE DETERMINATION

We considered this topic to become familiar with the principles and pitfalls of determining ground motion amplitudes from digital seismograms. There are several lessons worth remembering.

- In the general case of narrowband seismograph systems, the output of the recording system is rarely proportional to any physical units. The logical units for digital seismograms from such an instrument is counts. Only for very well-behaved broadband recording systems does it make sense to attach physical units directly to digital seismograms.
- We have learned how to quantitatively determine the different types of frequency response functions (magnification curves) which are used to determine ground motion amplitudes from digital seismograms and how these can be converted into each other.
- Any frequency response function is described by a frequency - dependent term $F(j\omega)$ which is assumed to be dimensionless and a frequency - independent scale factor $C$ whose units are determined by the output signal in whatever unit the system puts out (V, counts) divided by the units of the input signal for which the frequency response function is defined (displacement, velocity, acceleration).
- We have seen how the scale factor can be calculated in various ways and how it relates to calibration gain or channel sensitivity at a given calibration or normalization frequency.
- We have learned how to set up the corresponding instrument calibration files in GSE format for use with PITSA.
- Finally, we have seen how ground motion amplitudes are commonly calculated from digital seismograms by dividing the signal amplitude in counts by the value of the amplitude of the displacement frequency response function (magnification curve) at the signal frequency. **However, we should keep in mind that this procedure is based on the very important assumption that the input signal is harmonic.** A better procedure for non-harmonic signals would involve the deconvolution of the instrument response within a certain frequency band and measuring the amplitude from the deconvolved trace.

## 10.2 The properties of seismic onsets

A major portion of classical seismogram analysis deals with the recognition and analysis of seismic onsets. Onset properties are used in numerous contexts ranging from earthquake location and the determination of focal mechanisms to source modelling and traveltime tomography. Quite recently, the properties of seismic onsets have received additional attention based on the fact that they may allow the distinction of different friction laws discussed in the context of modelling earthquake sources (Iio, 1992). While it is easy to describe onsets in qualitative terms such as impulsive, emergent, strong or weak, with few exceptions (e.g. Seidl and Stammler, 1984, Seidl and Hellweg, 1988) a quantitative treatment of onsets has been rarely used within observational seismology. We will follow the onset definition of Seidl and Stammler (1984) which is based on the order of discontinuity at the signal front of a causal signal $f(t)$ ($f(t) = 0$ for $t < 0$).

> Definition: The discontinuity of the signal front at $t = 0$ is called an onset of order $p$, if $f^{(p)}(0+)$ is the first non-zero derivative.

In this definition, the time $t = 0$ corresponds to the arrival time of the wavelet which is commonly referred to as the onset time while the sign of the discontinuity (+1 or -1) defines the onset polarity. The unit step function $u(t)$ ($u(t) = 0$ for $t < 0$ and $u(t) = 1$ for $t \geq 0$) corresponds to an onset of order zero. The signal $f(t) = C \cdot t^a \cdot e^{-(b \cdot t)} \cdot u(t)$ on the other hand corresponds to an onset of order $a$. A special case of this signal for $a = 1$ is the Brune model far field source time function (Brune, 1970, 1971), that has been used in Fig. 10.2. It corresponds to an onset of order 1.

Onset time and onset order can be quite severely affected by the filter properties of the recording system used. This is illustrated in detail in Fig. 10.11 which shows a blow up of Fig. 10.2, centered around the theoretical onset time for the synthetic ground displacement trace in the top trace.

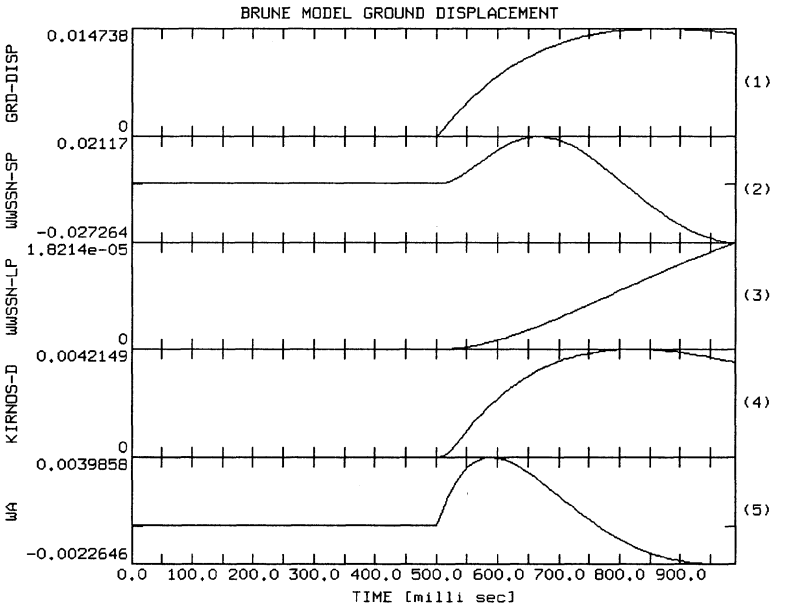



**Fig. 10.11** Onset distortion of a far field Brune model ground displacement pulse (onset order 1, top trace) by standard seismograph systems (traces 2-5). Displayed in traces 2 to trace 5 are the output signals of simulated WWSSN-SP, WWSSN-LP, KIRNOS, and a Wood-Anderson instrument for the synthetic input signal shown in trace 1.

Both changes in the steepness of the initial slope of the wavelets (onset order) as well as apparent time delays can be recognized for the onsets in traces 2-5. Again, imagine the determination of these parameters in a more realistic scenario with noisy traces.

### 10.2.1 ONSET ORDER

The potential change of the onset order by seismic recording systems is most conveniently understood from the initial value theorem of the Laplace transform for causal functions (Seidl and Stammler, 1984). In chapter 2.4, we have defined the transfer function using the definition of the bilateral or two-sided Laplace transform. In this case we obtained different impulse response functions for a given transfer function depending on the region of convergence (ROC). Most importantly, we learned why we may obtain acausal impulse response functions by using the inverse Fourier transform, or in the discrete case the inverse DFT. This might otherwise be quite confusing to anybody trying to numerically implement the calculation of impulse response functions from the transfer function. In the context of the digital anti-alias filter problem in chapter 8, we also saw that 'acausal' impulse response functions are very relevant to modern seismic acquisition systems.

On the other hand, most physical systems are causal and the impulse response function of the analog part of seismic acquisition systems (Fig. 6.19) will be a causal function as well. If the input signal to such a system is also causal, as in the case of seismic onsets, the introduction of the unilateral or one-sided Laplace transform has certain benefits. For example it will allow us to easily relate some of the distortions of seismic onsets to the frequency domain properties of the seismic recording system. It is worth mentioning that the causality of functions implies certain symmetries for the corresponding integral transforms (Fourier, Laplace, z). Causality for example places strong constraints on the types of attenuation laws permissible in nature (Aki and Richards, 1980). A detailed treatment of the mathematics of causality in a seismological context can be found in Ben-Menahem and Singh (1981, Appendix K).

*The unilateral Laplace transform*

The unilateral Laplace transform of a causal function $f(t)$ ($f(t) = 0$ for $t < 0$) is defined as:

$$\mathcal{L}_I[f(t)] = \int_{0+}^{\infty} f(t)e^{-st}\,dt \tag{10.21}$$

with the complex variable $s = \sigma + j\omega$. $\mathcal{L}_I[f(t)]$ will be written as $F_I(s)$ to distinguish it from the two-sided Laplace transform defined in equation (2.25). The lower integration limit is to be interpreted as approaching zero from positive times. The inverse transform, defined as

$$\mathcal{L}_I^{-1}[F_I(s)] = f(t) = \frac{1}{2\pi j}\int_{\sigma - j\infty}^{\sigma + j\infty} F_I(s)e^{st}ds \tag{10.22}$$

is formally equivalent to (2.42). The path of integration must lie in the region of convergence. Since the region of convergence for causal functions must be a right-sided half-plane, the region of convergence considered for the unilateral inverse transform always lies to the right of the right-most pole. While most of the properties of the unilateral and bilateral Laplace transform are equivalent, what makes the unilateral transform special in the present context is that the transform of a derivative incorporates initial values of the time function.

If $f(t)$ and $F_I(s)$ are a unilateral Laplace transform pair, the transform of the derivative $\dot{f}(t)$ becomes

$$\mathcal{L}_I[\dot{f}(t)] = sF_I(s) - f(0+) \tag{10.23}$$

with $f(0+)$ being the initial value of $f(t)$ ( $\lim_{t \to 0} f(t)$ with the limit taken for positive times). Similarly, for $n$ higher order derivatives, initial values of $f(t)$ and of its $n-1$ derivatives have to be incorporated

$$\mathcal{L}_I[f^{(n)}(t)] = s^n F_I(s) - \sum_{k=1}^{n} s^{n-k} \cdot f^{(k-1)} \cdot f(0+) \tag{10.24}$$

Proofs for (10.23) and (10.24) can be found in Kraniauskas (1992). With equation (10.23) the value of $\mathcal{L}_I[\dot{f}(t)]$ for $s \to \infty$ becomes

$$\lim_{s \to \infty} \mathcal{L}_I[\dot{f}(t)] = \lim_{s \to \infty} \int_{0+}^{\infty} \dot{f}(t) e^{-st}\, dt \tag{10.25}$$

If $\dot{f}(t)$ does not contain impulses at the origin (which by the sifting property of the delta function would always make the integral final), for causal functions only defined for $t > 0$, the value of the integral tends to zero for $s \to \infty$. Therefore

$$\lim_{s \to \infty} \int_{0+}^{\infty} \dot{f}(t) e^{-st}\, dt = \lim_{s \to \infty} sF_I(s) - f(0+) = 0 \tag{10.26}$$

or

$$f(0+) = \lim_{t \to 0} f(t) = \lim_{s \to \infty} sF_I(s) \tag{10.27}$$

Equation (10.27) is called the **initial value theorem** of the unilateral Laplace transform. Imagining $F_I(s)$ as the transfer function of a seismic recording system and $f(0+)$ as the shape of the impulse response function at $t = 0$ (onset shape), equation (10.27) states that the high frequency limit of the corresponding frequency response function (for $s \to j\omega$) will control the shape of a seismic onset ($f(t)$ for $t \to 0$ ).

For a seismic onset of order $p$, $f^{(p)}(0+) = f_p$ is the first non-zero derivative and $f^{(m)}(0+) = 0$ for $m < p$. Following the derivation of Seidl and Stammler (1984) we

combine the time derivative property (10.24) and the initial value theorem (10.27) to obtain

$$f^{(p)}(0+) = \lim_{t \to 0} f^{(p)}(t) = \lim_{s \to \infty} s^{p+1} F_I(s) = f_p \tag{10.28}$$

Dividing equation (10.28) by $s^{p+1}$ gives

$$\lim_{s \to \infty} F_I(s) = \frac{f_p}{s^{p+1}} \tag{10.29}$$

Since $1/s^{p+1}$ is the Laplace transform of $(t^p/p!) \cdot u(t)$ with $u(t)$ being the unit step function (Oppenheim and Willsky, 1983), the waveform of the onset is approximately

$$f(t) \approx \frac{f_p}{p!} \cdot t^p \text{ for } t = 0+ \text{ (Seidl and Stammler, 1984).} \tag{10.30}$$

In other words, if for $s \to \infty$ the transfer function of the recording system is proportional to $s^{-(p+1)}$ with $p$ an integer, the onset shape takes the particularly simple form of being proportional to $t^p$ for $t \approx 0+$.

Since the classical standardized seismometer-galvanometer systems (e.g. WWSSN-SP, WWSSN-LP, KIRNOS, Wood-Anderson) can all be effectively approximated by a rational transfer function of the form

$$T(s) = \frac{-\beta_L \cdot \prod_{k=1}^{L} (s - s_{0k})}{\alpha_N \cdot \prod_{k=1}^{N} (s - s_{pk})} = -\frac{\beta_0 + \beta_1 s + \beta_2 s^2 + \ldots + \beta_L s^L}{\alpha_0 + \alpha_1 s + \alpha_2 s^2 + \ldots + \alpha_N s^N} \tag{10.31}$$

with N poles at the roots of the denominator polynomial and L zeros at the roots of the numerator polynomial (Seidl, 1980), for $s \to \infty$ we obtain

$$\lim_{s \to \infty} T(s) \to -\frac{\beta_L s^L}{\alpha_N s^N} = c \cdot s^{L-N} \tag{10.32}$$

Therefore, for $s \to \infty$ the Laplace transform of a causal signal with onset order $p$ recorded by a seismometer system with transfer function $T(s)$ becomes

$$\lim_{s \to \infty} T(s)F_I(s) \approx c \cdot s^{L-N} \cdot \frac{f_p}{s^{p+1}} . \tag{10.33}$$

Using the same argument as for the derivation of equation (10.30) we obtain for the approximate wavelet shape at $t = 0+$:

$$f(t) \approx \frac{c \cdot f_p}{(L-N+p)!} \cdot t^{N-L+p} \tag{10.34}$$

Thus a recording system with a rational transfer function with N poles and L zeros reduces the order of an onset and hence changes the initial slope of the onset by $N - L$. This is the effect seen in Fig. 10.11.

### 10.2.2 ONSET POLARITY

Although the waveform shape and the initial slope of the onset were seen to be considerably altered by the different recording systems in Fig. 10.11, the sign of the onset remained the same for all the traces. This can be understood from equation (10.34) since the scaling factor $((c \cdot f_p)/(L-N+p)!)$ of the onset shape is independent of frequency. Therefore, there can be no frequency dependent change in the onset polarity for recording systems with rational transfer functions. The onset polarity is solely dependent upon the initial normalization of $c$ which is usually positive. The observation of different onset polarities in different frequency bands, which is not uncommon in seismological observatory practice, has its most likely explanation therefore in source processes or propagation effects.

### 10.2.3 FILTER DELAYS AND ONSET TIME

The determination of onset times is without doubt the most frequently performed task in classical seismological observatory practice. Since the solution of the earthquake location problem is crucial to essentially any seismological study, loosely speaking, onset time determinations are the 'bread and butter' of observational seismology. For this reason, the effects of filters on onset times deserve special attention. In chapter 8, we have already seen one way of dealing with the onset time effects caused by zero phase FIR filters such as are commonly used as digital anti-alias filters in modern seismic acquisition systems. From the initial value theorem of the Laplace transform and the discussion in the previous chapter we have learned that the wavelet shape at the onset is strongly influenced by the high frequency properties of the filters and that the initial slope of an onset

may be considerably changed by a particular recording system. In the case of a noisy signal, the noise may hide the true onset until it becomes strong enough to be recognized. This by itself may cause an apparent delay in the onset times determined from seismic records (Fig. 10.12).

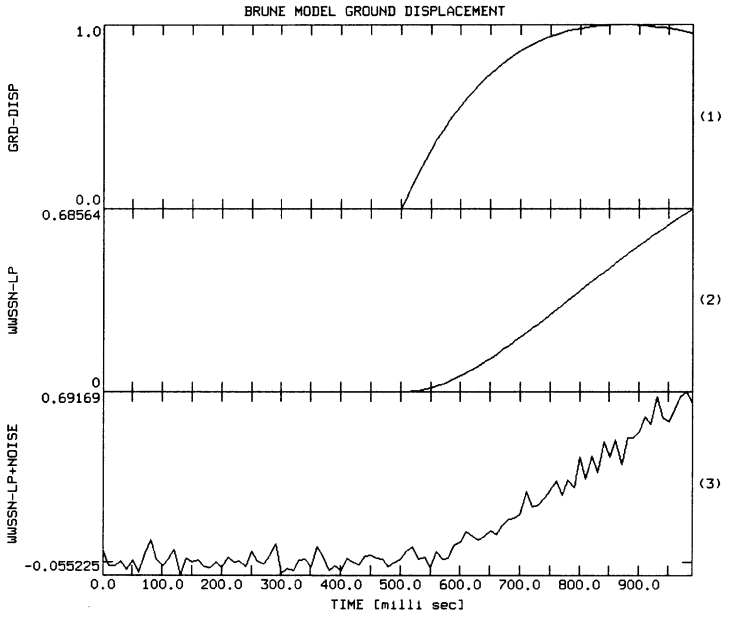



**Fig. 10.12** Apparent onset delay caused by the change of the onset order. The top trace shows a synthetic seismic source signal (Brune, 1970, 1971) which is filtered by a WWSSN-LP system (second trace).The result of superimposing Gaussian noise on the second trace is displayed in the bottom trace. Since the true onset is hidden in noise the apparent onset is delayed.

In addition, however, any filter with a non-zero phase function causes waveform distortion which may lead to measurable distortions of the onset times.

In chapter 2.3 (see also Problem 2.1) we have seen that a filter with the frequency response function $T(j\omega) = |T(j\omega)| \cdot e^{j\Phi(\omega)}$ changes a purely harmonic signal $A \cdot \sin(\omega_0 \cdot t)$ into $A \cdot |T(j\omega)| \cdot \sin(\omega_0 \cdot t + \Phi(\omega_0))$ . The phase $\Phi(\omega)$ gives the amount by which the phase of a harmonic signal with frequency $\omega$ is advanced [in radians].

This leads to the definition of the **phase delay** $t_{ph}(\omega)$ of a filter in [s] which is defined as

$$t_{ph}(\omega) = -\Phi(\omega)/\omega \ . \tag{10.35}$$

In other words, a harmonic signal with frequency $\omega$ which is filtered with a frequency response $|T(j\omega)| \cdot e^{j\Phi(\omega)}$ is delayed by $t_{ph}(\omega) = -\Phi(\omega)/\omega$ seconds. However, with the possible exception of some volcanic and industrial noise signals, seismic signals are only seldom purely harmonic.

In the case of arbitrary signals, the phase delay for a particular frequency only describes the delay of a single harmonic component. In the special case, however, that the phase response $\Phi(\omega)$ is a linear function of frequency, the shifting theorem of the Fourier transform (see chapter 2.3) states that the corresponding filter performs a constant time shift of the complete input signal. The frequency response function of a pure time-shifting filter has a modulus of one and a phase that is a linear function of frequency:

$$\text{Time-shifting filter: } T(j\omega) = |T(j\omega)| \cdot e^{j\Phi(\omega)} = e^{-j\omega a} \tag{10.36}$$

Here $a$ is the delay introduced by the time shifting filter in seconds.

In the case of general phase response functions, the time delays of harmonic signals are clearly defined by the phase delay. For signals other than purely harmonic the situation is more complicated. In the case of narrowband signals, the delay of the signal envelope (or center of gravity) is described by the **group delay** function $t_{gr}(\omega)$ :

$$t_{gr}(\omega) = -\frac{d}{d\omega}\Phi(\omega) \ . \tag{10.37}$$

For broadband signals, group delay and phase delay have no meaning. In this case the parameter of interest, is the delay of the onset, the **signal front delay** $t_{fr}$ defined as

$$t_{fr} = \lim_{\omega \to \infty} \frac{-\Phi(\omega)}{\omega} = \lim_{\omega \to \infty} t_{ph}(\omega) \ . \tag{10.38}$$

In the case of seismograph systems with rational transfer functions, the phase response function becomes constant for $\omega \to \infty$. Intuitively this can be understood from the method for graphically determining the frequency response function of a transfer function given by its poles and zeros in the complex s-plane (c.f. chapter 3.2 Graphical estimation of the frequency response function). To determine the phase part of the frequency

response function, the phase angles for the vectors from the zero locations in the s-plane to the point $j\omega$ on the imaginary axis were added together. Then, the phase angles of all the vectors from pole locations to the point $j\omega$ on the imaginary axis were subtracted. For $j\omega \to \infty$, because of the large distance between $j\omega$ and pole and zero locations, however, these angles no longer change with a change in frequency. This yields the important results that for analog seismograph systems which can be described by rational transfer functions

$$t_{fr} = \lim_{\omega \to \infty} \frac{-\Phi(\omega)}{\omega} = \lim_{\omega \to \infty} \frac{const}{\omega} = 0 \text{ (Seidl and Stammler, 1984).} \qquad (10.39)$$

At first glance this result suggests that in the context of onset time determination we only have to worry about apparent onset time delays caused by changes in the onset order of a seismic wavelet. Unfortunately, in the context of simulation filtering of digital seismic data there is more to it as we will see below.

*Delays and onset distortion in digital filters*

When simulating seismometer systems as FIR filters or, more generally, when evaluating inverse DFTs from discrete spectra, a type of onset distortion sometimes shows up which is similar in appearance to the effects discussed in chapter *8 The digital anti-alias filter*. We have seen an example in Problem 2.3 and in Fig. 8.9. It can be very confusing on first glance, since it shows up as acausality in the time sequence, although from the pole-zero positions of the corresponding transfer function, one may expect the system to be completely causal. So one may suspect incorrectly that something has gone wrong with the filtering.

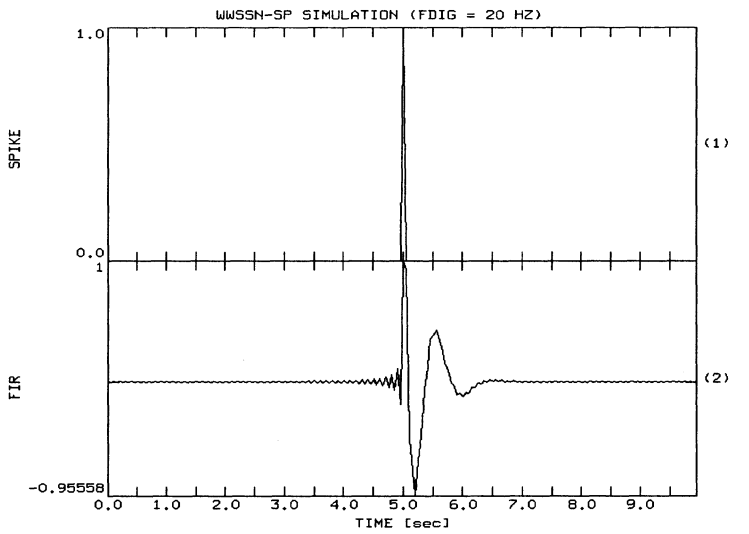



**Fig. 10.13** Impulse response function for a WWSSN SP instrument, calculated for a sampling frequency of 20 Hz (bottom trace). The input impulse is displayed in the top trace.

The effect is demonstrated in Fig. 10.13, which shows the impulse response for a WWSSN SP instrument, calculated for a sampling rate of 20 Hz. The sequence shown was calculated by inverse FFT of the frequency response function of the WWSSN SP instrument calculated from its poles and zeros, multiplied with the spectrum of the spike shown in the top trace. Note the oscillatory signals before the actual onset of the impulse response function.

It is important to note that as far as the pole-zero positions of the transfer function are concerned, the impulse response function should be completely causal. We are confronted with a truncation effect similar to the one discussed in chapter 9.2.2, this time in the frequency domain. When we think of a discrete impulse, we are actually considering a sequence of $N$ points of which $N-1$ are zero and one is non-zero. We tend to view this sequence as the result of sampling an impulsive time signal with sampling rate $\Delta T$. However, the continuous impulse has a Fourier transform of 1 for all frequencies, while the discrete impulse has a DFT (after proper scaling) of 1 for frequencies from 0 to the Nyquist frequency $f_{Nyquist} = 1/(2\Delta T)$. Due to the sampling theorem, the discrete impulse does not contain spectral energy above the Nyquist frequency. Hence, we can view the spectrum of the discrete spike as corresponding to the spectrum of the continuous spike (which is 1) multiplied with a boxcar window in the frequency domain and sampled at equally spaced frequencies. The inverse of the boxcar function in the frequency domain, however, is not an impulse but $\sin(\omega_c t)/(2\pi\omega_c t)$, a highly oscillatory function (Fig. 10.14).

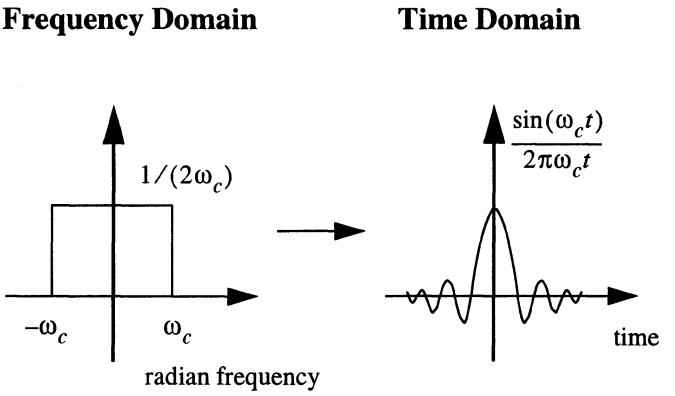



**Fig. 10.14** Boxcar tapering window in frequency and time domain.

How does this fit together? When we calculate the discrete inverse transform from a spectrum which is 1 for frequencies between 0 and $f_{Nyquist}$, should we not obtain a single pulse at the first sample? No! Since in this case $\omega_c = 2\pi f_{Nyquist}$, the oscillations of $\sin(\omega_c t)/(2\pi\omega_c t)$ occur with the Nyquist frequency. The reason why we do not see the oscillations is that all the samples of $\sin(\omega_c t)/(2\pi\omega_c t)$ - except for the first sample - are taken exactly at zero crossings. This situation changes if we apply a time shift to the signal which is not an integer multiple of $\Delta T$. According to the shifting theorem, this can be done in the frequency domain by applying a linear phase shift of the appropriate amount to the spectrum. This effect is demonstrated in Fig. 10.15.

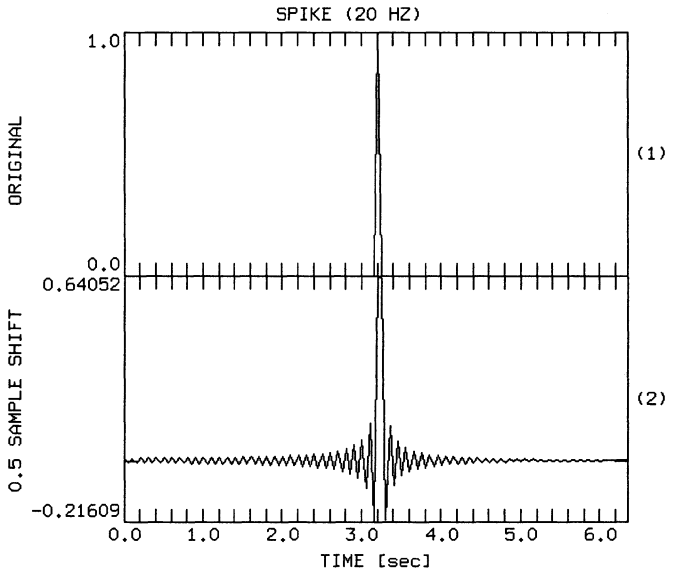



**Fig. 10.15** Shifting a discrete impulse by a non-integer multiple of one sample. The lower trace was obtained by shifting the impulse in the upper trace by 0.5 samples.

The spurious acausal signal components seen in Fig. 10.13 are the effect of such a shift by a non-integer multiple of the sampling interval. This time shift, therefore, corresponds to the phase delay at the Nyquist frequency (modulo the sampling interval) which for the present example (sampling frequency of 20 Hz) is equivalent to 0.422 samples. If we correct the impulse response function shown in Fig. 10.13 for this shift, we obtain an impulse response function more or less clean of spurious acausal oscillations (Fig. 10.16). However, we have to keep in mind that in order to remove the acausality effect, we pay the price of introducing a constant time shift to the impulse response function. If we use the onset-corrected trace for simulation filtering using (9.10), we must also apply a correction to the time tag of the filtered trace.

*Problem 10.11* Can the implementation of the WWSSN SP seismometer by bilinear transform preserve the phase properties of the analog system? Argue from the fact that the resulting recursive filter has a causal impulse response function.

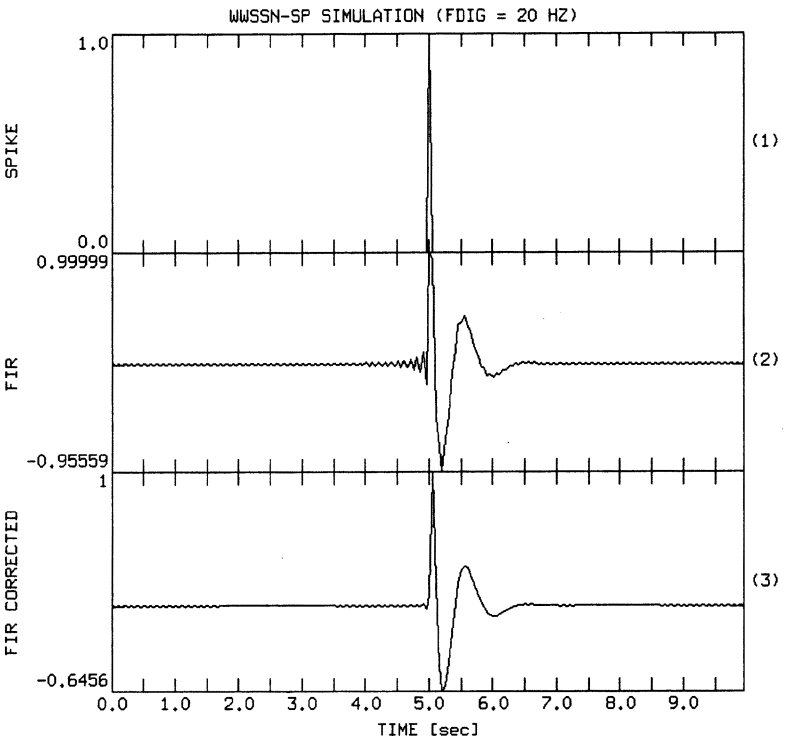



**Fig. 10.16** Onset delay correction for the impulse response function shown in Fig. 10.13. The second trace shows the uncorrected impulse response. The bottom trace shows the result of shifting the second trace by the negative phase delay time at the Nyquist frequency.

## 10.3 Rise time and pulse duration

It is commonly understood in earthquake seismology, that in addition to other effects such as attenuation or site effects, the duration of seismic signals observed at the earth's surface are dependent on the size of the earthquake source and the amount of stress release which has taken place. As a consequence, both the determination of source size and stress release are based on the measurement of characteristic seismogram times, either in the form of pulse duration, rise time, or in the spectral domain as spectral corner frequencies (Haskell, 1964; Aki, 1967, 1968; Okada, 1970; Brune, 1970, 1971; O'Neill and Healy, 1973; Boatwright, 1980, 1984). It is therefore essential to understand the limitations and constraints imposed on these parameters by the seismic recording system.

Qualitatively, the rise time of a signal is understood as the time a signal level raises from zero to some defined value, most commonly the maximum or final amplitude. In the con-

text of seismic pulse shapes, the pulse rise time is commonly defined as the interval between the intersection of the tangent of the steepest rise of a wavelet with the zero level and peak pulse amplitude (Fig. 10.17, e.g. Gladwin and Stacey, 1974)

$$t_r = a_{max}/slope_{max} \tag{10.40}$$

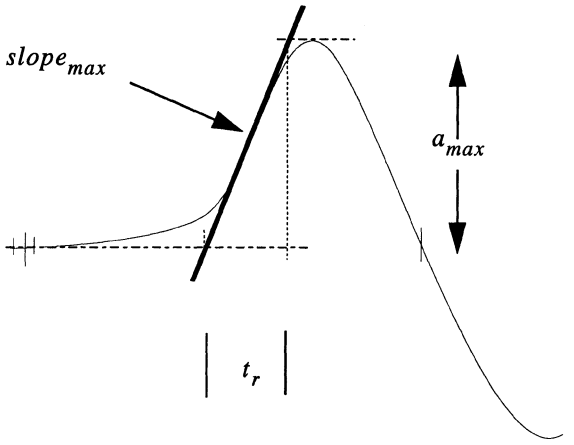



**Fig. 10.17** Definition of pulse rise time $t_r$.

*Problem 10.12* Calculate the rise time from the minimum to the maximum amplitude for a sinusoidal signal $A_0 \cdot \sin(\omega t)$.

If we approximate a seismic wavelet by ramps of constant slope, we can also see that in this case the rise time could also be measured as pulse duration (from zero crossings) in the differentiated trace (Fig. 10.18).

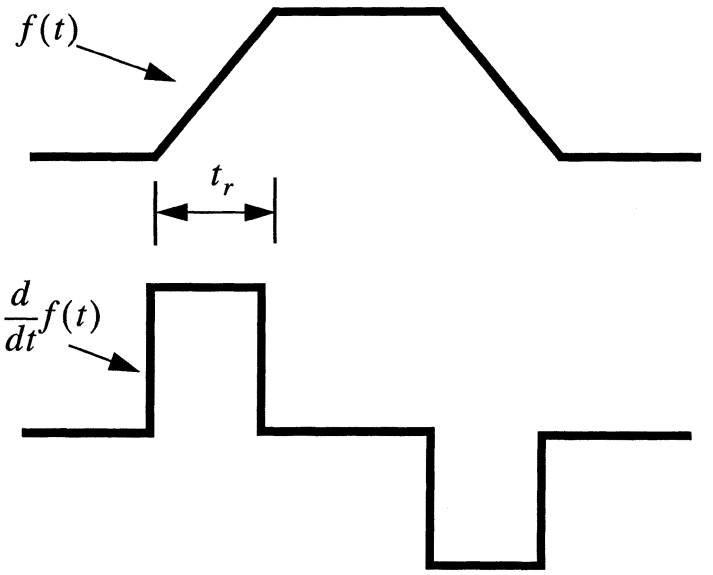



**Fig. 10.18** Rise time $t_r$ and pulse duration of the time derivative for ramp like wavelets.

In the case of more complicated broadband signals, rise time and pulse duration in both the original as well as the differentiated trace will still correlate, although not in the idealized way sketched above. It should also be noted that the definition of pulse duration as time between zero crossings is only one amongst others which are in use such as the equivalent width $t_e$, which is the width of the equal area rectangle, the energy spread width $\Delta t_s$, a width measure similar to the standard deviation in statistics, the mean square width $\overline{t^2}$, and the root mean square width $(\overline{t^2})^{1/2}$ (Bath, 1974). Each of these measures has its own merits and justifications, although in the present context these differences are not relevant. An extensive discussion can be found in Bath (1974). Here the fact is important that rise time and pulse duration measurements will in general correlate and that these parameters are affected by the seismic recording system in very similar ways. For this reason, the following discussion will focus on the rise time and assume that similar arguments can be made for any of the pulse duration definitions.

Seismic recording systems in general act as bandpass filters which limit the frequency band of the seismic signal. Since the rise time of a signal is inversely related to its dominant signal frequency, the longest rise time which theoretically could be measured from a seismic recording system is controlled by the eigenfrequency of the seismometer, while the shortest rise time is controlled by the lowpass filters involved (anti-alias filter, galva-

nometer). Because of numerous other sources of long period noise as well as model inadequacies (e.g. source complexities for large earthquakes), only the high frequency limitations are of practical importance for actual rise time determinations and will be discussed here. The effect of the low frequency limitations of the frequency band, however, are essential for the determination of the signal moment and will be treated in the following chapter.

To illustrate these limitations more quantitatively, the displacement step response functions for commonly used standard seismometer systems are displayed in Fig. 10.19. Even on casual inspection of Fig. 10.19 we see that the effects are quite severe.

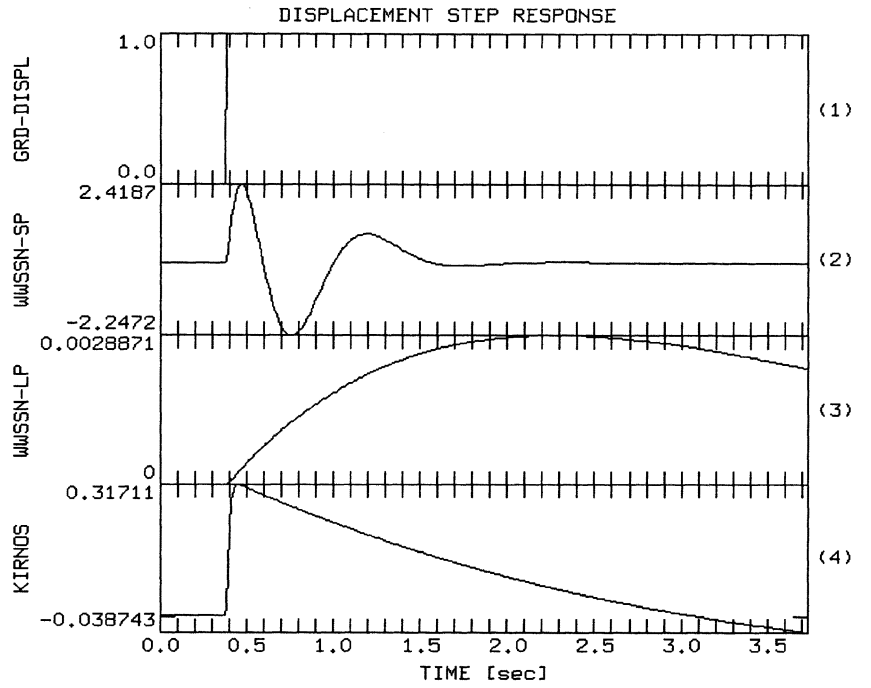



**Fig. 10.19** Simulated displacement step response functions for commonly used standard seismograph systems.

In earlier parts of this chapter we made use of the fact that the inverse Fourier transform of a boxcar function of amplitude $1/(2a)$ and width $2a$ in the frequency domain is the function $\sin(at)/(2\pi at)$ in the time domain (see Fig. 10.14). We will now exploit this fact to derive a quantitative expression for the rise time of the step response function for the case of an ideal lowpass filter.

Let's assume the anti-alias filter of a fictitious seismic recording system can be approximately described by an ideal zero phase lowpass filter which completely suppresses all frequencies above a corner frequency $f_c$. In this case, the amplitude frequency response of the filter would be described by a boxcar function in the frequency domain (Fig. 10.20).

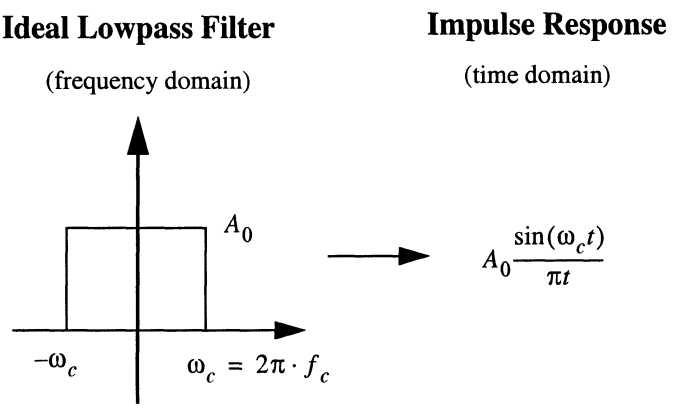



**Fig. 10.20** Ideal lowpass filter.

In this case, impulse and step response functions would be given by

$$A_0 \frac{\sin(\omega_c t)}{\pi t} \tag{10.41}$$

and

$$\frac{A_0}{2} + \frac{A_0}{\pi} \cdot \int_{-\infty}^{t} \frac{\sin(\omega_c \tau)}{\tau} d\tau = \frac{A_0}{2}\left(1 + \frac{2}{\pi} Si(\omega_c t)\right) \tag{10.42}$$

respectively (Papoulis, 1962). Here $Si(t)$ is the sine integral.

Evaluating these expressions for the maximum slope and the maximum amplitude yields according to (10.40) for the rise time

$$t_{r(idealLP)} = 1/(2 \cdot f_c) \text{ (Papoulis, 1962).} \tag{10.43}$$

For a zero phase lowpass filter with arbitrary spectral shape $A(\omega)$, the rise time becomes

$$t_r = \frac{A_0 2\pi}{\int_{-\infty}^{\infty} A(\omega)d\omega} \quad \text{(Papoulis, 1962)} \tag{10.44}$$

which can be interpreted as the rise time of an ideal lowpass filter with the same spectral amplitude at zero frequency and equivalent area of the amplitude spectrum. An example for impulse response and step response functions for an ideal lowpass filter with a rise time of 2 sec is shown in Fig. 10.21.

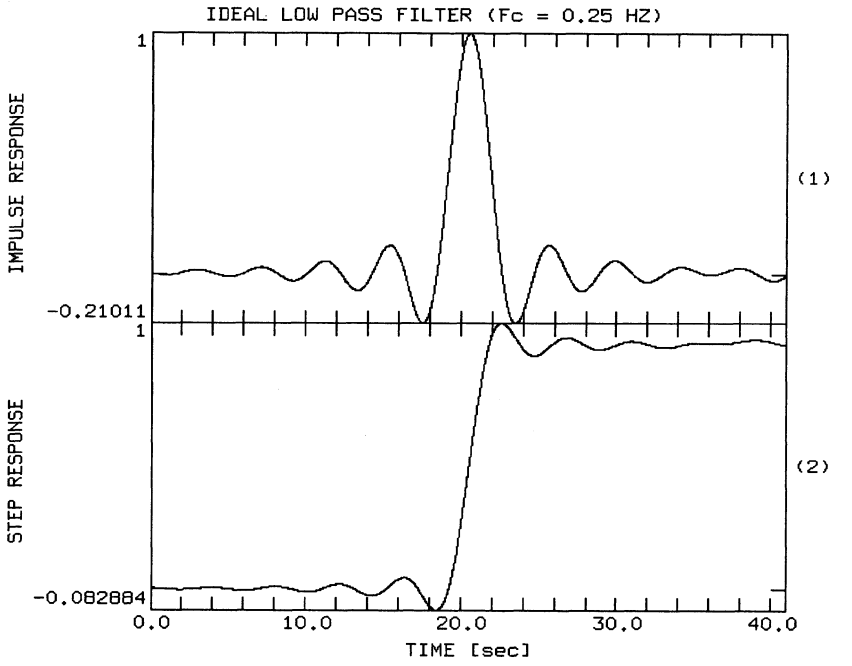



**Fig. 10.21** Impulse and step response function for an ideal (boxcar) lowpass filter with a rise time of 2 sec. The traces are shifted to the center of the plot for display purpose.

Note that expressions (10.43) and (10.44) are strictly valid only for zero phase lowpass filters. Nevertheless they provide an easy way to get a rough estimate of the minimum rise time which can be determined from a given recording system.

*Problem 10.13* Calculate the impulse and step response functions for a causal and an acausal Butterworth LP filter with 8 poles and a corner frequency of 1 Hz using DST. Discuss time delays and the minimum detectable pulse duration and rise times for potential filter signals.

## 10.4 Signal moment

One of the most basic physical parameters used to measure the earthquake strength is the seismic moment $M_0$ defined as

$$M_0 = \mu \cdot \bar{u} \cdot A \tag{10.45}$$

with $\mu$ the shear modulus, $\bar{u}$ the average slip, and $A$ the fault area (Aki and Richards, 1980). It is related to the spectral level of the farfield displacement pulse at frequency zero $U(0)$ by a constant factor $C$ containing the correction for radiation pattern, geometrical amplitude spreading, attenuation and the effects of the free surface (see also chapter 1)

$$M_0 = C \cdot U(0) \tag{10.46}$$

From the definition of the Fourier transform in equation (2.8) we can see that $U(0)$ can be calculated in either the frequency or the time domain since

$$U(0) = \int_{-\infty}^{\infty} u(t)e^{-j2\pi ft}\, dt = \int_{-\infty}^{\infty} u(t)e^{-j2\pi 0t}\, dt = \int_{-\infty}^{\infty} u(t)\, dt = m_s \tag{10.47}$$

Hence, the spectral level at frequency zero $U(0)$ equals the pulse area in the time domain, a parameter commonly referred to as signal moment $m_s$ (e.g. Seidl and Hellweg, 1988). This relationship is demonstrated in Fig. 10.22 for a synthetic Brune source pulse (Brune, 1970, 1971). The top trace has been scaled to a signal moment of 1 which can be seen to be equal to the low frequency spectral level in the bottom trace.

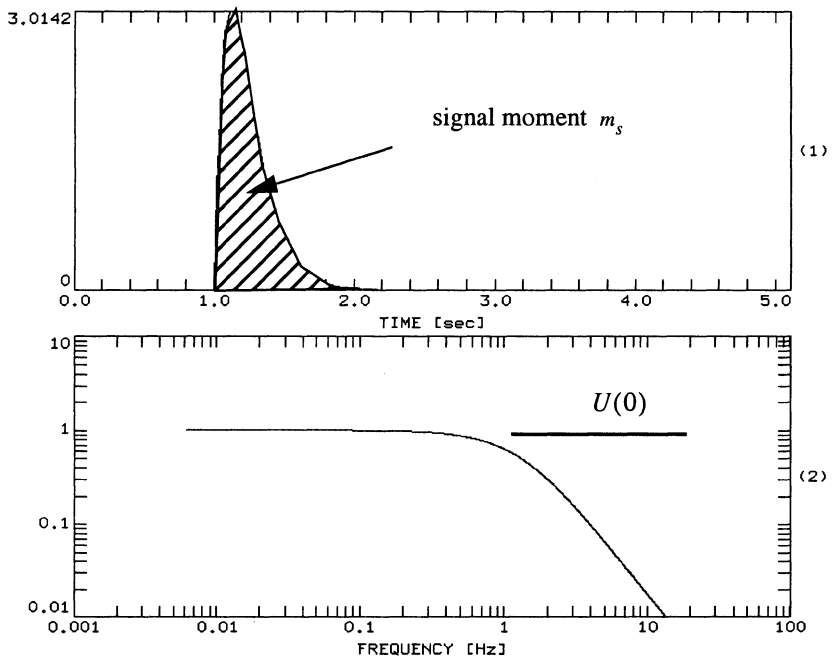



**Fig. 10.22** The determination of signal moment $m_s$ and $U(0)$ .

Fig. 10.22, however, illustrates a very unrealistic situation in the sense that the traces have been simulated without the effects of a seismometer. A farfield displacement spectrum as recorded on an instrument with frequency response function $T(j\omega)$ will be

$$S(j\omega) = U(j\omega) \cdot T(j\omega) \tag{10.48}$$

Since pendulum seismometers act as highpass filters $T(j\omega) = 0$ for $\omega = 0$. Therefore, the signal moment of the corresponding seismogram $s(t)$ becomes zero. This is illustrated in Fig. 10.23.

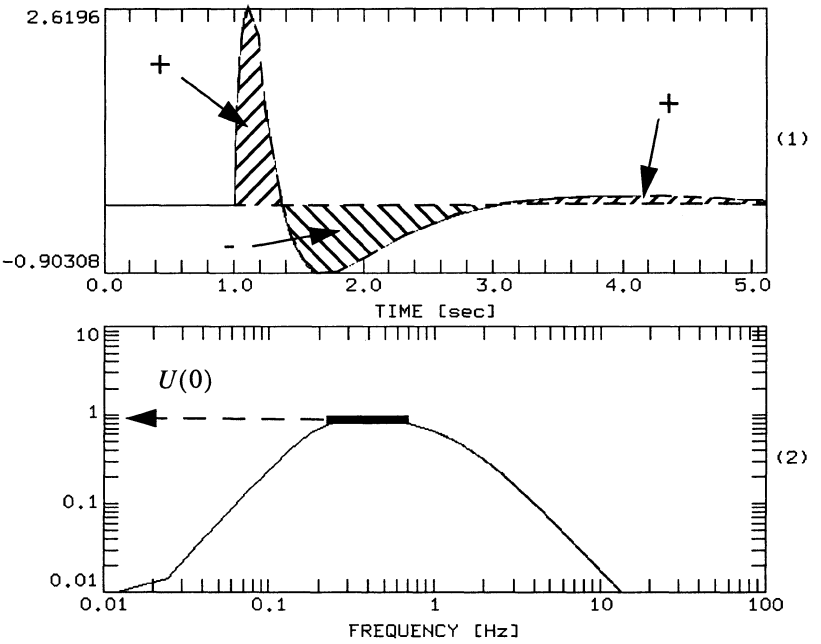



**Fig. 10.23** Signal moment $m_s$ and $U(0)$ for the signal from Fig. 10.22 as it would be recorded on a 5 sec displacement seismometer with a damping factor of 0.707.

This causes the displacement signal to become two-sided so that the positive and negative shaded areas cancel. As a consequence, the seismogram signal moment as it would be determined for a finite time integral between the onset and the first zero crossing (displacement integral method) will underestimate the signal moment of the actual ground motion. For a quantitative discussion of the corresponding errors - which depend on both the pulse duration and the seismometer period - and potential methods to correct for them see Seidl and Hellweg (1988).

In the frequency domain, we usually deal with this problem by assuming that the low frequency plateau values seen in the band-limited signal can be extrapolated down to zero frequency as is illustrated in the bottom trace of Fig. 10.23.

# References

Aki, K., Scaling law of seismic spectrums, J. Geophys. Res., 72, 1217-1231,1967.

Aki, K., Seismic displacements near a fault, J. Geophys. Res., 73, 5359-5376, 1968.

Aki., K., and P. Richards, Quantitative seismology, Freeman Publ., 1980.

Bath, M., Spectral analysis in geophysics, Elsevier Scientific Publishing Company, 1974.

Ben-Menahem, A., and S. J. Singh, Seismic waves and sources, Springer Verlag New York, 1981.

Boatwright, J., A spectral theory for circular sources; Simple estimates of source dimension, dynamic stress drop, and radiated seismic energy, Bull. Seismol. Soc. Am., 70, 1-26, 1980.

Boatwright, J., Seismic estimates of stress release, J. Geophys. Res., 89, 6961-6968, 1984.

Brune, J. N., Tectonic stress and the spectra of seismic shear waves from earthquakes, J. Geophys. Res., 75, 4997-5009, 1970.

Brune, J.N., Correction to 'Tectonic stress and the spectra of seismic shear waves from earthquakes', J. Geophys. Res., 76, 5002, 1971.

Dornboos, D. J., E. R. Engdahl, and T. H. Jordan, International Seismological Observing Period, Preliminary Science Plan, USGS, Box 25046, MS 967, Denver Federal Center, Denver, CO 80225, 1990.

Ellsworth, W. L.,and G. C. Beroza, Seismic evidence for an earthquake nucleation phase, Science, 268, 851-855, 1995.

Fowler, J. C., The effects of non causal filters in the PASSCAL instrument, IRIS internal report, 1992.

Gladwin, M.T., and F.D. Stacey, Anelastic degradation of acoustic pulses in rock, Phys. Earth and Plan. Int., 8, 332-336, 1974.

GSE Wave Form Data Format, Version CRP190/rev.3: Ad hoc group of Scientific Experts, Geneva, Annex D2, D4-D10, D48-D62, 1990.

GSE Wave Form Data Format, Version CRP243: Ad hoc group of Scientific Experts, Geneva, OpsAnnex 3, 1995.

Gurrola, H., G. E. Baker, and J. B. Minster, Simultaneous time-domain deconvolution with application to the computation of receiver functions, Geophys. J. Int., 120, 537-543, 1995.

Hamming, R. W., Digital filters, Sec. Edition, Prentice-Hall Signal Processing Series, A.V. Oppenheim (series editor), 1977.

Haskell, N. A., Total energy and energy spectral density of elastic wave radiation from propagating faults, Bull. Seismol. Soc. Am., 54, 1811-1841, 1964.

Hjelt, S. E, Pragmatic inversion of geophysical data, Lecture Notes in Earth Sciences 39, Springer-Verlag, 1992.

Iio, Y, Slow initial phase of the P wave velocity pulse generated by microearthquakes, GRL, 19, 477-480, 1992.

IEEE Digital Signal Processing Committee, Programs for Digital Signal Processing. IEEE Press, John Wiley & Sons, Inc. New York, 1979.

Iio, Y., Observations of the slow initial phase generated by microearthquakes: Implications for earthquake nucleation and propagation, J. Geophys. Res., 100, 15333-15349, 1995.

Jaeger, R.C., 1982., Tutorial: Analog data aquisition technology, part II - Analog to Digital Conversion, IEEE Micro, Vol. 2, 46 - 56, August 1982.

Kraniauskas, P., Transforms in signals and systems, Addison-Wesley, 1992.

Kulhanek, O., Anatomy of seismograms, Elsevier, 1990.

Luzitano, R. D., Data report for dense array recordings from nine aftershocks of the July 21, 1986 Earthquake in Chalfant Valley, California, USGS open file report 88-71, 1988.

McClellan, J. H., T. W. Parks, and L. R. Rabiner, FIR Linear phase filter design program. Programs for digital Signal Processing, IEEE Press New York, 1979.

Menke, W., Geophysical Data Analysis: Discrete Inverse Theory, Academic Press, 1984.

Mitra, S. K., and J. F. Kaiser (Ed.), Handbook for digital signal processing, J. Wiley & Sons, 1993.

O'Dowd, R. J., Ill-conditioning and pre-whitening in seismic deconvolution, Geophys. J. Int., 101, 489-491, 1990.

Okada, Y., The difference in the durations of P and S waves from propagating faults, J. Phys. Earth, 18, 193-202, 1970.

Oldenburg, D.W., A comprehensive solution to the linear deconvolution problem, Geophys. J. R. astr. Soc., 65, 331-357, 1981.

O'Neill, M. E., and J. H. Healy, Determination of source parameters of small earthquakes from P-wave rise time, Bull. Seismol. Soc. Am., 63, 599-614, 1973.

Oppenheim, A. V., and R. W. Schafer, Discrete-Time Signal Processing, Prentice Hall, 1989.

Oppenheim, A. V., and A. S. Willsky, Signals and systems, Prentice-Hall, Inc., 1983.

Papoulis, A., The Fourier integral and its application, Mc Graw Hill, 1962.

Park, S., Principles of Sigma-Delta Modulation for Analog-to-Digital Converters, Motorola, APR8/D Rev. 1, October 1999 (http://mot-sps.com/books/apnotes/pdf/APR8.pdf)

Parks T. W., and C. S. Burrus, Digital Filter Design, J. Wiley & Sons,1987.

Plesinger, A., Fast inversion of transient calibration responses of arbitrary digital seismograph systems, Computers & Geosciences, 19, 189-201, 1993.

Plesinger, A., M. Zmeskal, and J. Zednik, PREPROC - Software for automated preprocessing of digital seismic data, Version 2.1, Ed. E. Bergman, NEIC Golden/GI Prague, April 1995.

Press, W. H., S. A. Teukolsky, W.T. Vetterling, and B. P. Flannery, Numerical Recipes in C, Second Edition, Cambridge University Press, 1992.

Proakis, J. G., and D. G. Manolakis, Digital Signal Processing Principles, Algorithms, and Applications, Macmillan Publ. Company, 1992.

Proceedings of the seminar on deconvolution of seismograms and high-fidelity seismometry, J. Geophys, 39, 501-626, 1973.

Quatieri, T. F., and J. M. Tribolet, Computation of the real cepstrum and minimum-phase reconstruction. Programs for Digital Signal Processing. IEEE Press, John Wiley & Sons, Inc. New York, 1979.

Quanterra technical note, Linear and minimum-phase filters, Quanterra Inc., 1992.

Scherbaum, F., Short Course on First Principles of Digital Signal Processing for Seismologists, IASPEI Software Library, Vol. 5, 1993.

Scherbaum, F., and J. Johnson, PITSA, Programmable Interactive Toolbox for Seismological Analysis, IASPEI Software Library, Vol. 5, 1993.

Scherbaum, F., Basic Concepts in Digital Signal Processing for Seismologists, Lecture Notes in Earth Sciences 53, Springer Verlag, 1994.

Scherbaum, F. Zero Phase FIR filters in digital seismic acquisition systems: Blessing or Curse, EOS, 78, No. 33, 343-344, 1997.

Scherbaum, F., and M. P. Bouin, FIR Filter Effects and Nucleation phases, Geophys. J. Int., 130, 661-668, 1997.

SEED Reference Manual, IRIS, 1990.

Seidl, D. The simulation problem for broad-band seismograms, J. Geophys., 48, 84-93, 1980.

Seidl, D. and M. Hellweg, Restoration of broad-band seismograms (Part II): Signal moment determination, J. Geophys., 62, 158-162, 1988.

Seidl, D., and W. Stammler, Restoration of broad-band seismograms (Part I), J. Geophys., 54, 114-122, 1984.

Sipkin, A. A., and A. L. Lerner-Lam, Pulse-shape distortion introduced by broadband deconvolution, Bull. Seismol. Soc. Am., 82, 238-258, 1992.

Stearns, S. D., Digital Signal Analysis, Hayden Book Company, 1975.

Steim, J., The very-broad-band seismograph. Ph.D. thesis, Harvard University, Cambridge Massachusetts, 1986.

Stefánsson, R., R. Bödvarsson, R. Slunga, P. Einarsson, S. Jakobsdóttir, H. Bungum, S. Gregersen, J. Havskov, J. Helme, and H. Korhonen, Earthquake prediction research in the South Iceland Seismic Zone and the SIL project, Bull. Seismol. Soc. Am., 83, 696-716, 1993.

Strum, R. D., and D. E. Kirk, First principles of discrete systems and digital signal processing, Addison-Wesley, 1988.

Tarantola, A., Inverse Problem Theory, Elsevier, 1987.

Treitel, S., and L. R. Lines, Linear inverse theory and deconvolution, Geophysics, 47, 1153-1159, 1982.

Willmore, P. L. (Ed.), Manual of seismological observatory practice, WDCA report 20, U.S. Dept. of Commerce, NOAA, Boulder/CO, 1979.

Wielandt, E. , Elimination of the noncausal filter response from digital seismic records, unpubl. notes, 1992.

Wolfram, S., Mathematica, A system for doing mathematics by computer, Addison Wesley 1991.

Zmeskal, M. and Plesinger, A., Simulation of standard seismograms, in PREPROC - Software for automated preprocessing of digital seismic data, Version 2.1, Ed. E. Bergman, NEIC Golden/GI Prague, April 1995.

# Appendix: Solution to problems

## Chapter 2

*Solution 2.1* Using Euler's formula, we can write the sinusoidal signal as the sum of two exponential signals

$$A_i \sin(\omega_0 t) = A_i \cdot \frac{e^{j\omega_0 t} - e^{j(-\omega_0)t}}{2j} \tag{2.1-1}$$

Since we are dealing with a linear system, the output of the sum of the two input signals (here $A_i \cdot \frac{e^{j\omega_0 t}}{2j}$ and $-A_i \cdot \frac{e^{j(-\omega_0)t}}{2j}$) equals the sum of the individual output signals. These are

$A_i \frac{e^{j\omega_0 t}}{2j} \cdot |T(j\omega_0)| \cdot e^{j\Phi(\omega_0)}$ and $-A_i \frac{e^{j(-\omega_0)t}}{2j} \cdot |T(j(-\omega_0))| \cdot e^{j\Phi(-\omega_0)}$, respectively.

Therefore, the complete filter output signal is equal to:

$$A_i \frac{(|T(j\omega_0)| \cdot e^{j\Phi(\omega_0)}) e^{j\omega_0 t} - (|T(j(-\omega_0))| \cdot e^{j\Phi(-\omega_0)}) e^{j(-\omega_0)t}}{2j} \tag{2.1-2}$$

Because $|T(j(-\omega_0))| = |T(j\omega_0)|$ for the RC-filter (and for any real filter), we obtain

$$A_i \frac{|T(j\omega_0)|}{2j} (e^{j\omega_0 t + j\Phi(\omega_0)} - e^{j(-\omega_0)t + j\Phi(-\omega_0)}) \tag{2.1-3}$$

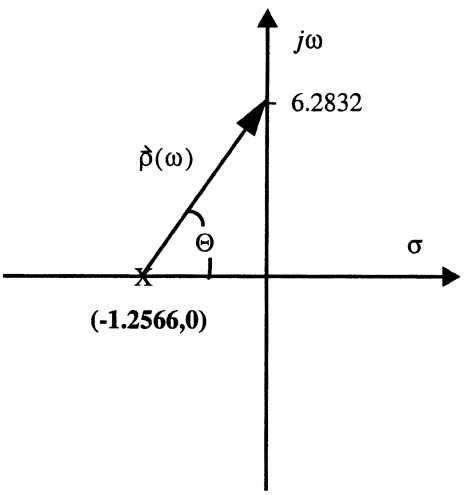
Since $\Phi(\omega_0) = \text{atan}(Im(T(j\omega_0))/Re(T(j\omega_0)))$ and $\text{atan}(-\alpha) = -\text{atan}(\alpha)$ we obtain $\Phi(-\omega_0) = -\Phi(\omega_0)$ and the output signal becomes

$$A_i \cdot \frac{|T(j\omega_0)|}{2j}\left(e^{j(\omega_0 t + \Phi(\omega_0))} - e^{-j(\omega_0 t + \Phi(\omega_0))}\right) = A_i \cdot |T(j\omega_0)| \cdot \sin(\omega_0 t + \Phi(\omega_0)) \tag{2.1-4}$$

Using equations (2.21) and (2.22) we obtain for the RC filter output

$$A_i \cdot \frac{1}{\sqrt{1 + (RC\omega_0)^2}} \sin(\omega_0 t - \text{atan}(\omega_0 RC)) \tag{2.1-5}$$

*Solution 2.2* The pole position is at $-1/\tau$. $\tau = R \cdot C = 4.0\Omega \cdot 0.1989495F$ which is $4.0V/A \cdot 0.1989495A\,\text{sec}/V = 0.795798\text{sec}$. Hence, the pole is at $-1.2566$ (rad/s). For each point on the imaginary axis (angular frequency axis), determine the reciprocal of the length of the vector from the pole to that point. You can do this either by using a ruler and graph paper or simply by exploiting analytical geometry. Plot this value as a function of angular frequency or frequency, respectively. Below, the procedure is demonstrated schematically for a frequency of 1Hz (Fig. A 2.1; note, that Fig. A 2.1 is not on 1:1 scale).




**Example: f=1 Hz**

$$\omega = 2\pi \cdot 1 = 6.2832$$

$$|\hat{p}(2\pi)| = 6.4076$$

$$|T(j2\pi)| = \frac{1}{\tau}\left[\frac{1}{|\hat{p}(\omega)|}\right]$$

$$|T(j2\pi)| = \frac{1.2566}{6.4076} = 0.1961$$

**Fig. A 2.1** Graphical determination of the modulus of the frequency response function for the RC filter of Problem 2.2. The plot demonstrates the evaluation for a frequency of 1 Hz.

*Solution 2.3* The Digital Seismology Tutor (DST) simulates the action of systems defined by their transfer function in the complex s-plane. As we will see later in more detail, a transfer function of a more general system can have more than one pole as well as a number of zeros (at which the transfer function becomes zero). The positions of poles and zeros define the transfer function completely. In order to do the filtering, DST needs to know the position(s) of the pole(s) and zero(s) (which will be introduced in later chapters) in the complex s-plane.

After starting up the DST, select the *Modify and Enter* option from the *Poles/Zeros* menu. In order to enter the pole position for the RC-filter at (-1.2566, 0), enter the real part ( -1.2566) into the uppermost left box and the imaginary part (0.0) into the uppermost right box as shown in Fig. A 2.2. Enter a 1.0 for the scale factor in the bottom box. Finally accept the input either by using the *OK* or the *Apply* button. The difference between these two is merely that the *OK* button closes the window after accepting the input while the *Apply* button leaves is open for further input.

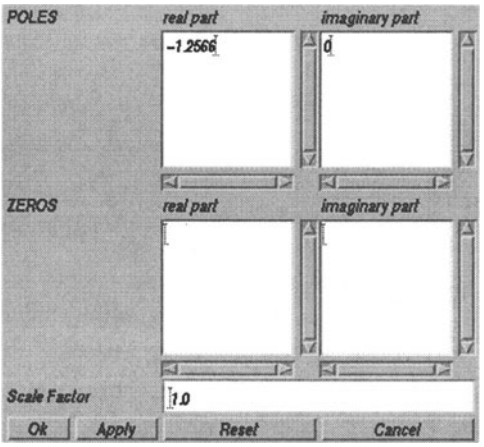



**Fig. A 2.2** How to enter the pole position for Problem 2.3 into the *Modify and Enter Poles/Zero* panel of the DST.

After you entered the pole into the DST, you can visualize its position within the complex s-plane by using the *Map* option from the *Poles/Zeros* menu (Fig. A 2.3).

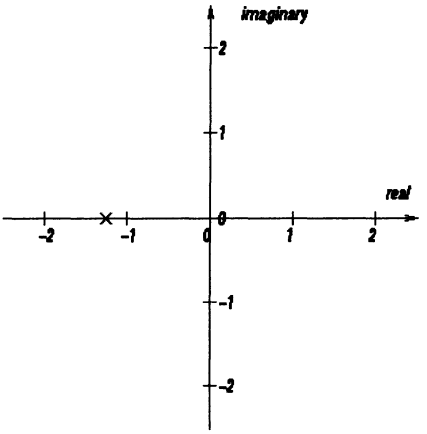



**Fig. A 2.3** Pole position of RC-filter from Problem 2.3 within the complex s-plane as calculated with DST.

The corresponding frequency response can be viewed using the *Show Response* menu from the DST main window using the option *Frequency Response -> Amplitude-> lin-lin* (Fig. A 2.4).

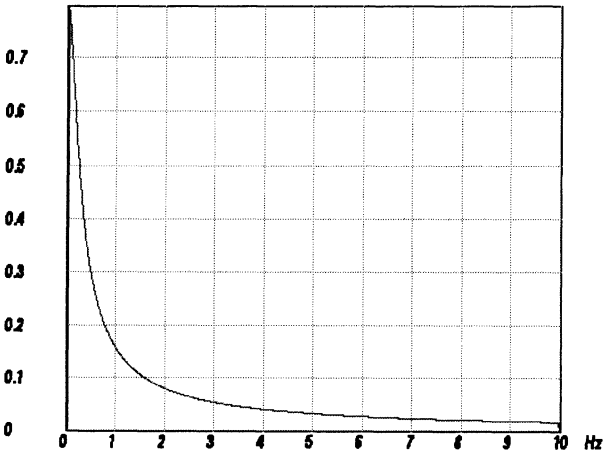



**Fig. A 2.4** Frequency response function (amplitude only) for the pole position shown in Fig. A 2.3.

**Note:** Although the plot in Fig. A 2.4 shows a continuous curve, it is actually a discrete approximation of the continuous frequency response function. The details of the underlying relationship will be explained in detail in chapter 7 "From infinitely continuous to

finite discrete". At this point it is sufficient to know that the internal sampling frequency for the calculation of the frequency response function can be modified in the *Setup* menu. For reasons which are explained in chapter 7, the frequency band which is plotted ranges from 0 to 1/2 of the internal sampling frequency.

*Solution 2.4* The balance at time $t = nT + T$:

$$y(nT+T) = y(nT) + \alpha y(nT) + \alpha x(nT) + x(nT) \tag{2.4-1}$$

$$= y(nT) + \alpha y(nT) + (\alpha+1)\cdot x(nT) \tag{2.4-2}$$

$$y(nT+T) - y(nT) = \alpha y(nT) + (\alpha+1)\cdot x(nT) \text{ (divide by T)} \tag{2.4-3}$$

$$\frac{y(nT+T)-y(nT)}{T} = \frac{\alpha}{T} y(nT) + \frac{(\alpha+1)}{T}\cdot x(nT) \tag{2.4-4}$$

Next, replace the difference quotient by the derivative and $nT$ by $t$:

$$\dot{y}(t) - \frac{\alpha}{T} y(t) = \frac{(\alpha+1)}{T}\cdot x(t) \tag{2.4-5}$$

Taking the Laplace transform yields

$$s\cdot Y(s) - \frac{\alpha}{T}\cdot Y(s) = \frac{\alpha+1}{T}\cdot X(s) \tag{2.4-6}$$

$$\left(s - \frac{\alpha}{T}\right)\cdot Y(s) = \frac{\alpha+1}{T}\cdot X(s) \tag{2.4-7}$$

$$T(s) = \frac{Y(s)}{X(s)} = \frac{\alpha+1}{T}\cdot\frac{1}{s-\frac{\alpha}{T}} \tag{2.4-8}$$

From the equation above we see that a pole exists at $s = \alpha/T$

The corresponding impulse response becomes: $\frac{\alpha+1}{T}\cdot e^{\frac{\alpha}{T}\cdot t}$

From the positive exponent we see that the impulse response is unstable! While this is normally an unwanted feature, for a checking account it is just was we expect!

## Chapter 3

*Solution 3.1* In order to obtain an internal sampling frequency of 100 Hz and a window length of 2048 points as recommended in Problem 3.1, these values have to be entered into the *Setup* menu (*Modify* option). For the spike position of the impulse response enter a value of 1024 to obtain a centered spike. To finally accept the input, you have to use either the *OK* or the *Accept* button. For part a) select the *Modify and Enter* option from the *Poles/Zeros* menu and enter the real and imaginary parts of each pole position on a separate line of the uppermost input boxes as shown in Fig. A 3.1.

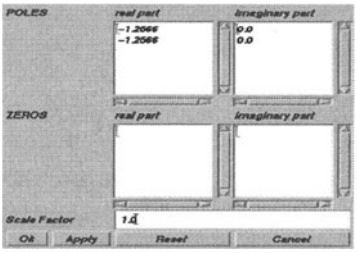



**Fig. A 3.1** Pole/Zero input panel for Problem 3.1a).

The corresponding impulse response function can be viewed using the *Impulse Response* option from the *Show Response* menu of the DST (Fig. A 3.2).

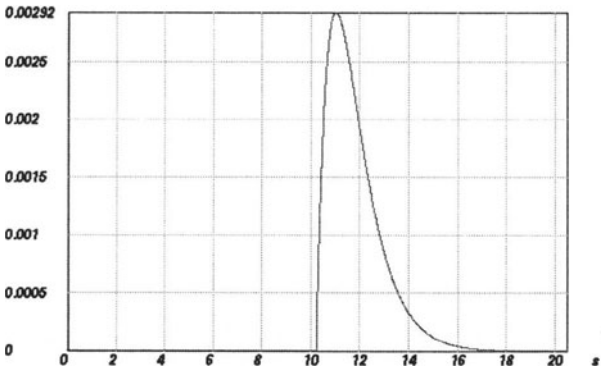



**Fig. A 3.2** Impulse response function for Problem 3.1a).

For Problem 3.1b), one pole falls into the left half of the s-plane while the other is located in the rigth half. The input for the corresponding pole positions is shown in Fig. A 3.3.

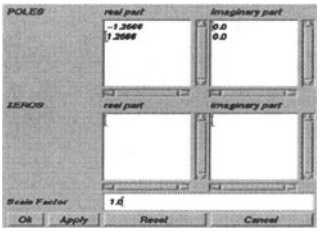



**Fig. A 3.3** Pole position input panel for Problem 3.1b).

Fig. A 3.4 displays the corresponding pole map (option *Map* from the *Poles/Zeros* menu).

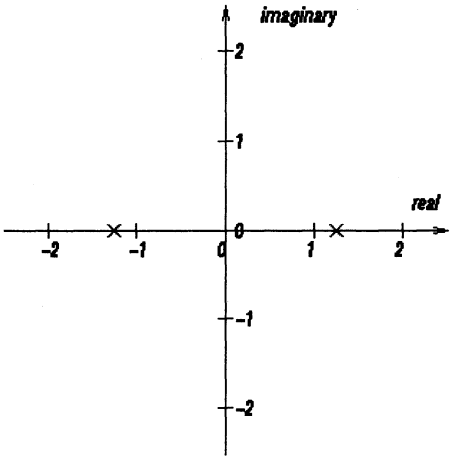



**Fig. A 3.4** Pole map for Problem 3.1b).

The resulting impulse response shows a signal symmetric to the position of the input spike Fig. A 3.5.

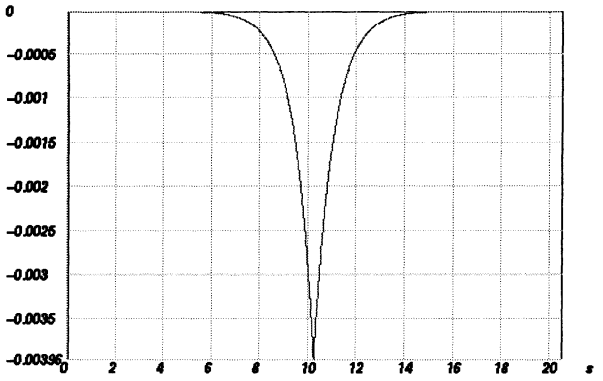



**Fig. A 3.5** Impulse response function for Problem 3.1b).

Finally, if the two poles are both put into the right half s-plane - Problem 3.1c) - the resulting impulse response function is a left-sided signal as shown in Fig. A 3.6.

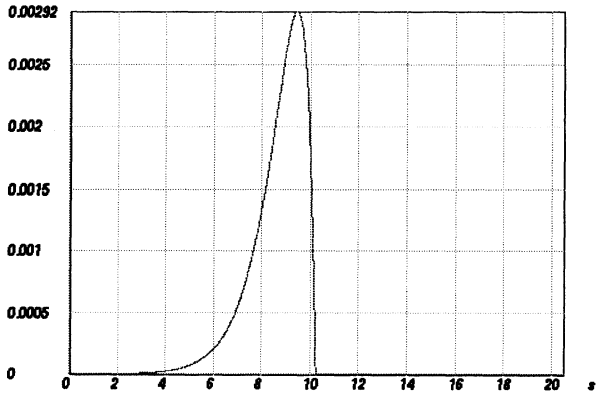



**Fig. A 3.6** Impulse response function for Problem 3.1c).

Hence, the resulting impulse response function will be a) right-sided, b) two-sided, and c) left sided which can be understood from the arguments in chapter 3.1. The amplitude portion of the frequency response function is identical in all three cases while the phase response differs.

With respect to Problem 2.3, the slope of the amplitude response is steeper by 20dB/decade since an additional pole is present.

*Solution 3.2* Since the pole and the zero become symmetrical to the imaginary axis, the pole vectors and zero vectors are always of equal length. Hence their ratio is always constant. It follows that the amplitude response is the same for all frequencies, characteristic of an allpass filter. Fig. A 3.7 shows the pole/zero distribution (a) and the corresponding impulse response function (b).

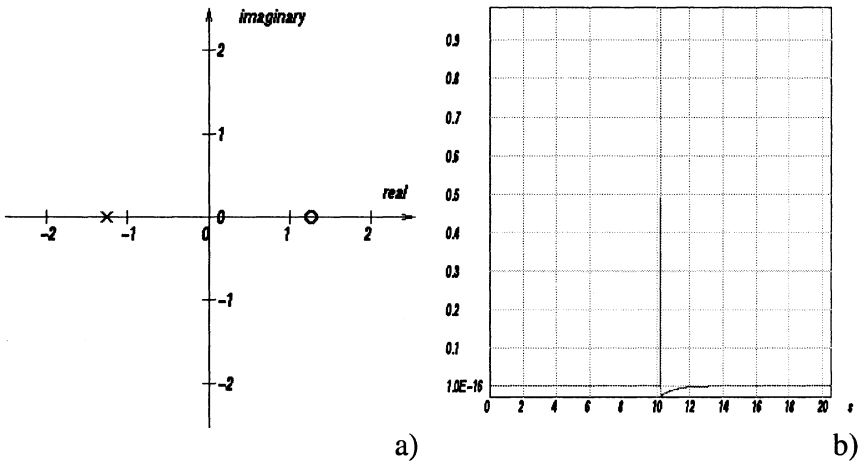



**Fig. A 3.7** a) Pole/zero distribution and b) input impulse function of the allpass filter of Problem 3.2. The amplitude part of the frequency reponse is constant for all frequencies (not shown).

*Solution 3.3* a) A general transfer function F(s) will be described by poles and zeros. Taking the inverse, 1/F(s) causes all poles to become zeros and zeros to become poles. For a minimum phase system, all the poles and zeros of the transfer function F(s) are on the left half s-plane. This will not change for the inverse system. Hence, the inverse system will also be minimum phase and therefore stable. b) The poles and zeros of a general mixed phase system with zeros on either half-plane can always be expanded into two systems. The allpass is constructed by taking all right half-plane zeros and adding symmetric poles. The minimum phase system is constructed by taking all poles (which have to be in the left half- plane anyway for stability reasons) and adding zeros where the allpass system had added poles for symmetry reasons. On multiplication of the two transfer functions these additional poles and zeros will cancel (Fig. A 3.8).

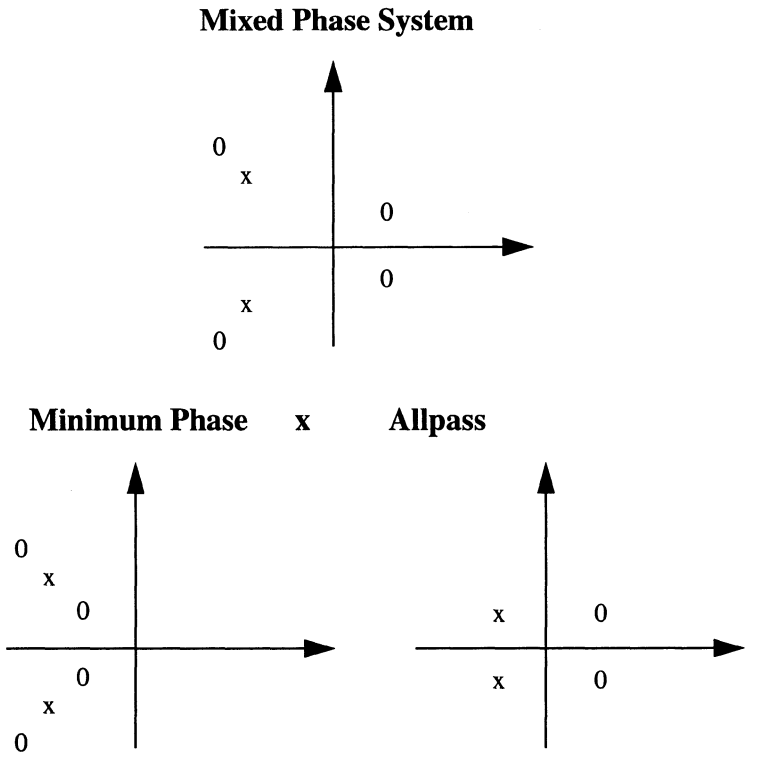



**Fig. A 3.8** Separating a mixed phase system into a minimum phase system and an allpass.

*Solution 3.4* Since the transfer function for Problem 3.1b has a pole on either side of the imaginary axis, we need to cancel the right-sided pole by a zero. In order not to change the amplitude characteristic, we have to do this with an allpass filter whose poles and zeros are symmetrical to the imaginary axis. The filter we need must have a zero at (+1.2566, 0) and a pole at (-1.2566, 0), which is exactly the allpass from Problem 3.2. The poles and zeros of the filter have to be added to the poles and zeros of the original system. Therefore, the filtered system will have two poles at (-1.2566, 0), one at (1.2566, 0), and a zero at (1.2566, 0) as shown in Fig. A 3.9.

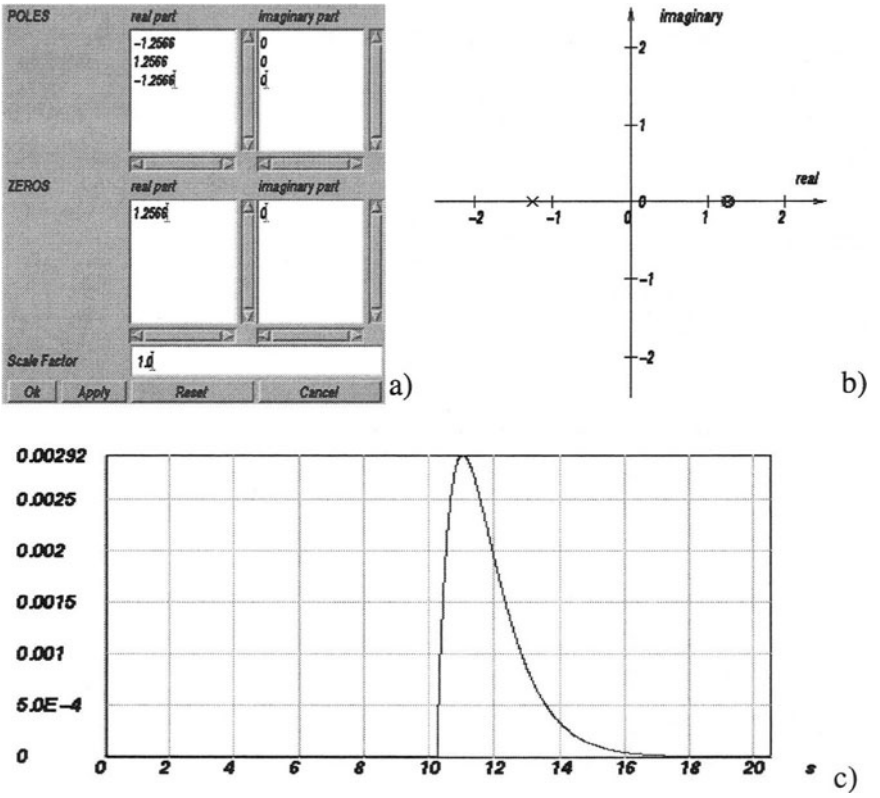



**Fig. A 3.9** *Poles and Zeros* input menu in DST (a), corresponding poles and zeros map (b), and impulse response function of the filtered system of Problem 3.1b (c). The internal sampling frequency in DST was set to 100 Hz (*Setup* menu).

*Solution 3.5* For this exercise, an internal sampling frequency of 20 Hz is recommended (*Modify* option in the *Setup* menu) in DST. First, select the *Modify and Enter* option from the *Poles/Zeros* menu and enter the real and imaginary parts of the pole and zero position. Fig. A 3.10 displays the pole/zero map and the amplitude part of the corresponding frequency response function.

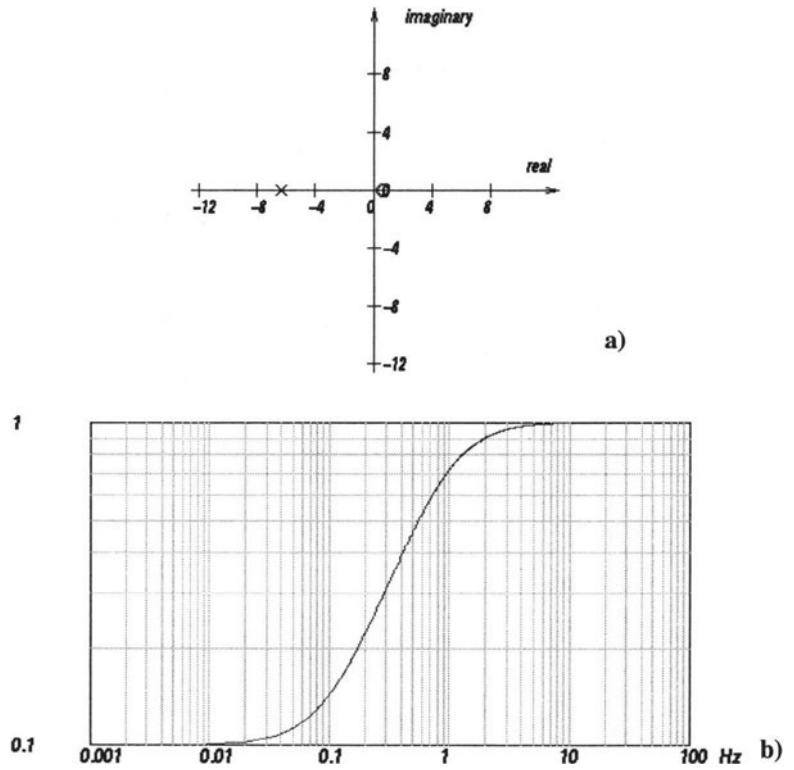



**Fig. A 3.10** Pole/zero map (a) and amplitude part of the frequency response function (b) for a system consisting of a pole at (-6.28318, 0.0) and a zero at (0.628318, 0.0). Fig. A 3.10 was created using DST with an internal sampling frequency of 20 Hz and 2048 points for the FFT (*Setup* menu).

From the pole and zero positions we expect corner frequencies at $f_1 = 0.1$ Hz and at $f_2 = 1$ Hz, respectively. We can see that the frequency response function starts out constant, increases by 20 db/decade at $f_1$ then decreases by 20 dB/decade at $f_2$.

*Solution 3.6* Start out with the double pole at (-1.2566, 0). For the successive steps, the pole positions would be $p_{1,2} = -1.2566(\cos\alpha + j\sin\alpha)$ with $\alpha$ being incremented from $15^o$ to $75^o$ in steps of $15^o$ (plus $\alpha = 85^o$, $\alpha$ is measured clockwise from the negative real s-axis).

The corresponding pole positions are given below:

| [ $^o$ ] | pole position |
|---|---|
| $0^o$ | $-1.2566 \pm j0.0$ |
| $15^o$ | $-1.21378 \pm j0.325232$ |
| $30^o$ | $-1.08825 \pm j0.6283$ |
| $45^o$ | $-0.88855 \pm j0.88855$ |
| $60^o$ | $-0.6283 \pm j1.08825$ |
| $75^o$ | $-0.325232 \pm j1.21378$ |
| $85^o$ | $-0.10952 \pm j1.25182$ |

Below, the resulting impulse response functions are shown for 30°, 60°, and 85° (Fig. A 3.11 - Fig. A 3.13).

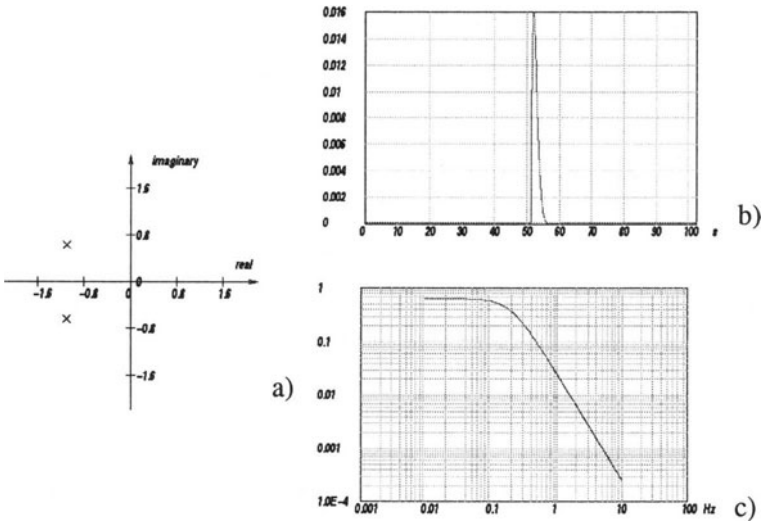



**Fig. A 3.11** Pole positions (a), impulse response (b) and frequency response (c) for a system with a conjugate complex pole pair with the pole positions at an angle of 30° from the real axis of the s-plane. The internal sampling frequency in DST was 20 Hz, the window length 2048 points, and the spike position for the impulse response1024 (*Setup* menu).

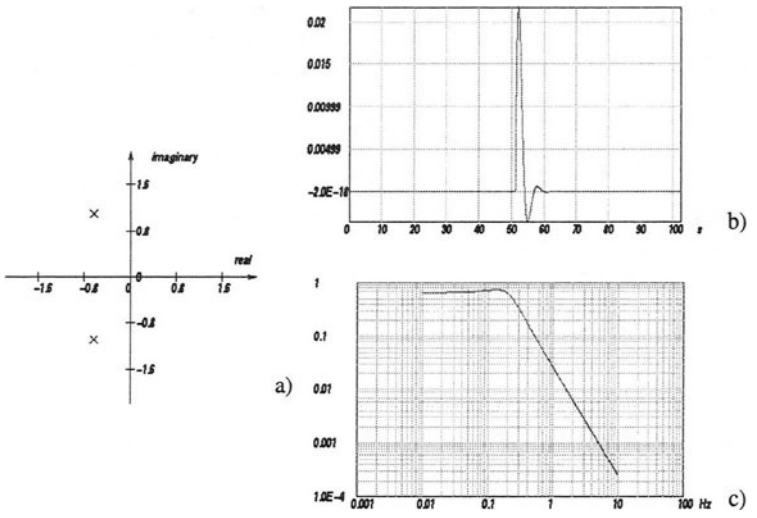



**Fig. A 3.12** Pole positions (a), impulse response (b), and frequency response (c) for a system with a conjugate complex pole pair with the pole positions at an angle of 60° from the real axis of the s-plane.

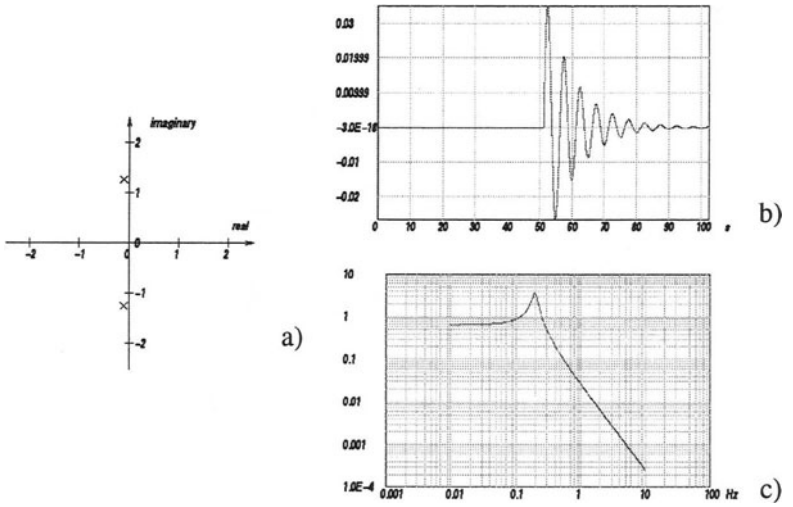



**Fig. A 3.13** Pole positions (a), impulse response (b), and frequency response (c) for a system with a conjugate complex pole pair with the pole positions at an angle of 85° from the real axis of the s-plane.

As the poles get closer to the imaginary axis, there are more oscillations in the impulse response. On the other hand when the poles are farther from the imaginary axis, the oscillations are more strongly damped. Notice that the spectrum changes significantly only near the corner frequency, not at the high and low ends of the spectrum.

*Solution 3.7* In Problem 3.6 we have seen that the amplitude part of the frequency response functions shows a strong peak if the poles were close to the imaginary axis. Hence, zeros close to the imaginary axis will correspond to strong selective suppression which is what we need. Therefore, we need to put zeros on or very close to the imaginary axis. In order to sharpen the notch, we also have to put poles close to the zeros, cancelling their effect for frequencies away from the notch frequency. One solution is shown below (Fig. A 3.14).

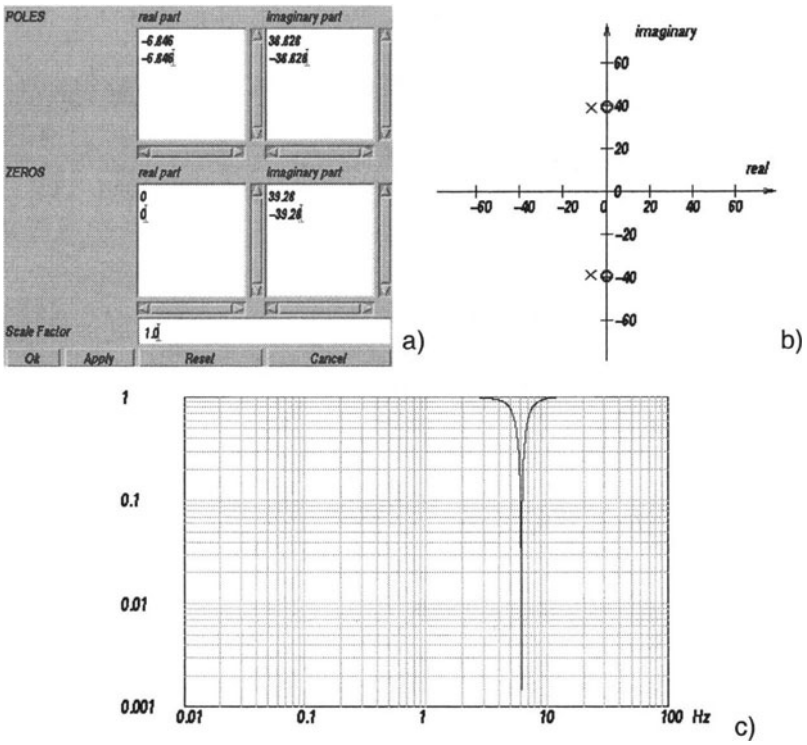



**Fig. A 3.14** *Poles and Zeros* input menu in DST (a), corresponding poles/zeros map (b), and amplitude frequency response function for a notch filter at 6.25 Hz.The corner frequency corresponding to the chosen pole position was 6.275 Hz, the corner frequency corresponding to the the chosen zero positions 6.25 Hz.The internal sampling frequency in DST was 20 Hz, the window length 2048 points, and the spike position for the impulse response 1024 (*Setup* menu).

*Solution 3.8* Within DST, the frequency response function of Problem 3.8 can be modeled by poles and zeros entered via the *Modify and Enter* option of the *Poles/Zeros* menu. First select *problem3.8* from the *Load Response to Fit* option of the *Poles/Zeros* menu to display the frequency response function shown in Fig. 3.4. Next, estimate the different slopes and determine the number of poles and zeros which are needed to model them. One reasonable interpretation is sketched in Fig. A 3.15. Next, try to find the corner frequencies at which the changes in slope occur (here 0.05 Hz and 5 Hz).

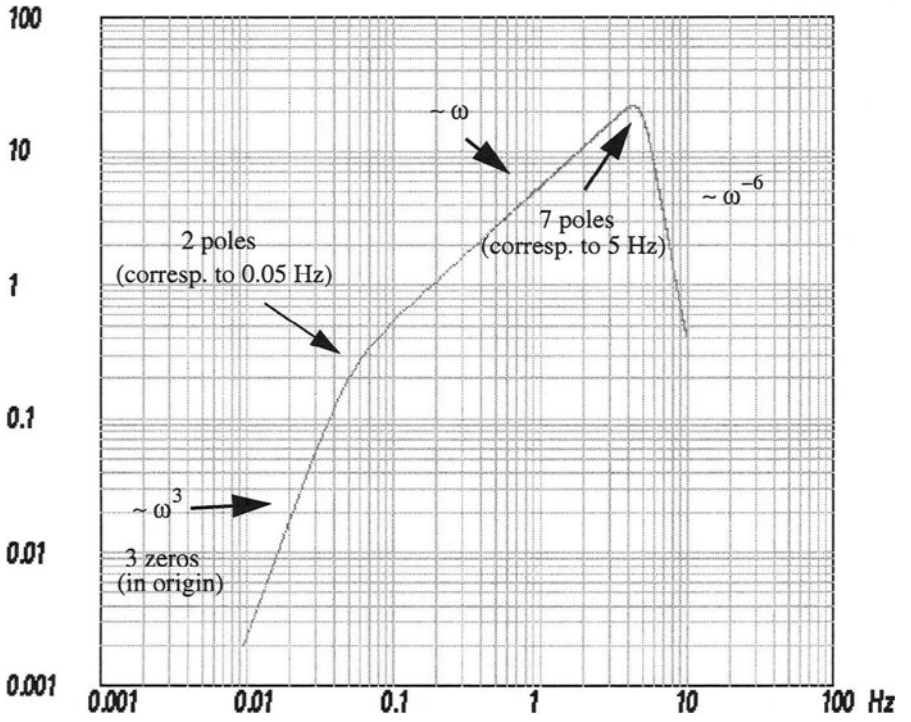



**Fig. A 3.15** Frequency response function (amplitude) with an 'unknown' pole - zero distribution from Problem 3.8.

Poles corresponding to the first corner frequency at 0.05 Hz have to be located either on the negative real axis of the s-plane at a distance of 0.31416 from the origin or occur in pairs defined by $p_{1,2}^{0.5Hz} = -0.31416 \cdot (\cos\alpha \pm j\sin\alpha)$ with $\alpha$ between 0 and 90°. If we choose 45° as a trial value, we obtain -0.2221 for the real part and ± 0.2221 for the imaginary part.

Poles corresponding to the second corner frequency at 5 Hz on the other hand have to be located either on the negative real axis at a distance of 31.416 from the origin or occur in pairs defined by $p_{1,2}^{5Hz} = -31.416 \cdot (\cos\alpha \pm j\sin\alpha)$ with $\alpha$ between 0 and 90°. If we put 1 pole on the negative real axis at (-31.416, 0) we are left with six poles to distribute as complex conjugate pairs. If we arbitrarily choose $\alpha$ to be 25, 50, and 75°, we obtain the following pole positions:

(-0.2221, 0.2221)
(-0.2221, -0.2221)
(-28.473, 13.277)
(-28.473, -13.277)
(-20.194, 24.066)
(-20.194, -24.066)
(-8.131, 30.346)
(-8.131, -30.346)
(-31.416, 0)

This is displayed below (Fig. A 3.16).

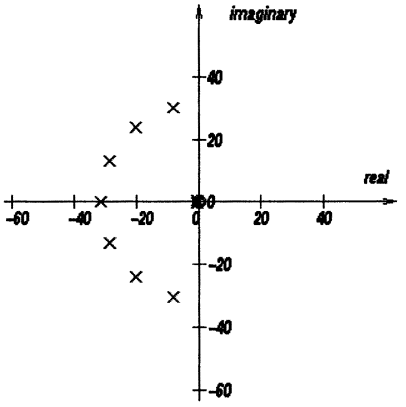



**Fig. A 3.16** Trial pole/zero position to match the amplitude frequency response function of Problem 3.8.

The corresponding amplitude frequency response function (not shown) matches the shape of the frequency response function shown in Fig. A 3.15 fairly well visually. The amplitude values, however, are too low by an approximate factor of $2.5 \cdot 10^{10}$. This can be adjusted by entering a scale factor of *2.5e10* in the *poles and zeros* input panel of DST.

Hint: For practical reasons regarding DST it is advisable to start out with the poles close to the origin and add the more distant poles only after the shape of the frequency response function has been matched in the low frequency range.

The resulting frequency response function now matches very closely the one given for Problem 3.8 (Fig. A 3.17).

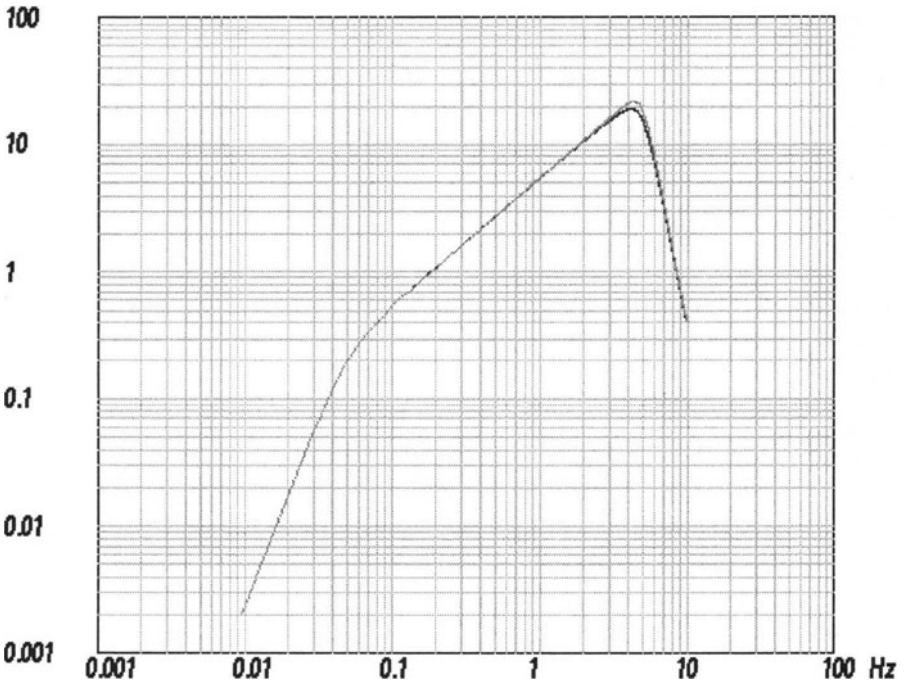



**Fig. A 3.17** Frequency response function of Problem 3.8 (upper curve) and frequency response function generated from the poles and zeros shown in Fig. A 3.16 with a scaling factor of 2.5e10.

At the corner frequency of 5 Hz, the frequency response function of Problem 3.8 has slightly higher amplitudes and shows a sharper change of slopes with respect to the trial frequency response function. This difference can be reduced by shifting some of the poles of the trial response closer to the imaginary axis to obtain a more "resonant" behaviour at that frequency.

The real distribution of poles and zeros for the frequency response function of Problem 3.8 is given below. It describes the frequency response function of the GRF array in SE Germany.

Poles:
(-0.2221, 0.2221)
(-0.2221, -0.2221)
(-7.0058, 30.6248)
(-7.0058, -30.6248)
(-19.5721, 24.5742)
(-19.5721, -24.5742)
(-28.3058, 13.6288)
(-28.3058, -13.6288)
(-31.4159, 0.0)
Zeros:
(0.0, 0.0)
(0.0, 0.0)
(0.0, 0.0)
Scale factor:
2.49059e10

For the generation of Fig. A 3.15 - Fig. A 3.17 within DST, an internal sampling frequency of 20 Hz and a window length of 2048 points was used.

## Chapter 4

*Solution 4.1* The time derivative of (4.26) is

$$\begin{aligned}\frac{dx_r(t)}{dt} = & -\varepsilon x_{r0} e^{-\varepsilon t}(c_1 \sin(\sqrt{\omega_0^2 - \varepsilon^2}t) + c_2 \cos(\sqrt{\omega_0^2 - \varepsilon^2}t)) \\ & + x_{r0} e^{-\varepsilon t} c_1 \sqrt{\omega_0^2 - \varepsilon^2} \cos(\sqrt{\omega_0^2 - \varepsilon^2}t) \\ & - x_{r0} e^{-\varepsilon t} c_2 \sqrt{\omega_0^2 - \varepsilon^2} \sin(\sqrt{\omega_0^2 - \varepsilon^2}t)\end{aligned} \tag{4.1-1}$$

With

$$c_3 = -(c_1\varepsilon + c_2\sqrt{\omega_0^2 - \varepsilon^2}) \tag{4.1-2}$$

and

$$c_4 = -(c_2\varepsilon - c_1\sqrt{\omega_0^2 - \varepsilon^2}) \tag{4.1-3}$$

we obtain

$$\frac{dx_r(t)}{dt} = x_{r0}e^{-\varepsilon t}(c_3\sin(\sqrt{\omega_0^2 - \varepsilon^2}t) + c_4\cos(\sqrt{\omega_0^2 - \varepsilon^2}t)) \tag{4.1-4}$$

As for the displacement seismometer, the amplitude ratio of two consecutive maxima or minima are solely determined by the exponential term. Therefore, equations (4.27) and (4.28) are valid for the determination of the damping constant of seismometers with velocity transducers as well.

*Solution 4.2* From Fig. 4.4 we can directly read the values of the first two consecutive peak amplitudes which are: $a_1 = 0.0869349$, the value of the maximum and $a_2 = |-0.014175| = 0.014175$, the value of the minimum of the calibration signal. Therefore, we obtain

$$\left(\frac{a_1}{a_2} = 6.13297\right) \Rightarrow \frac{\Lambda}{2} = 1.81368 \tag{4.2-1}$$

From $\Lambda = 3.6276$ we obtain h = 0.5. From Fig. 4.4 we measure the period from the second zero-crossing as $T \approx 1.15\,\text{sec}$ (the exact value would be 1.15470054). This yields $f_0 = 1Hz$.

*Solution 4.3* The displacement impulse response is the inverse Fourier transform of:

$$T(j\omega) = \frac{Output(j\omega)}{Input_{disp}(j\omega)} \tag{4.3-1}$$

Now let us examine the input signal, a step function in acceleration, which is equivalent to the integral of a spike. Hence, by equivalence between integration in the time domain

and division by $j\omega$ in the frequency domain, the frequency response of the input signal in acceleration is:

$$Input_{acc}(j\omega) = \frac{1}{j\omega} \cdot 1 \tag{4.3-2}$$

with 1 being the frequency response of an impulse (here in acceleration). In order to obtain the input signal in displacement, the input signal in acceleration has to be integrated twice. This corresponds to two times division by $j\omega$:

$$Input_{disp}(j\omega) = \left(\frac{1}{j\omega}\right)^2 \cdot Input_{acc}(j\omega) = \left(\frac{1}{j\omega}\right)^3 \tag{4.3-3}$$

Therefore,

$$T(j\omega) = \frac{Output_{disp}(j\omega)}{Input_{disp}(j\omega)} = 1 \Big/ \left(\frac{1}{j\omega}\right)^3 \tag{4.3-4}$$

which becomes

$$T(j\omega) = (j\omega)^3 \cdot Output_{vel}(j\omega) \tag{4.3-5}$$

In the time domain, multiplication with $(j\omega)^3$ corresponds to triple differentiation. Since high frequency noise will be greatly enhanced by differentiation, this explains why this approach to obtaining the impulse response is often impractical.

*Solution 4.4* The pole positions for the three different damping factors are:

$$h = 0.25: -1.5708 \pm j\,6.0837$$

$$h = 0.5: -3.14159 \pm j\,5.4414$$

$$h = 0.62: -3.89557 \pm j\,4.9298$$

The corresponding pole/zero maps, impulse responses and amplitude parts of the frequency response functions for these damping values are shown in Fig. A 4.1.

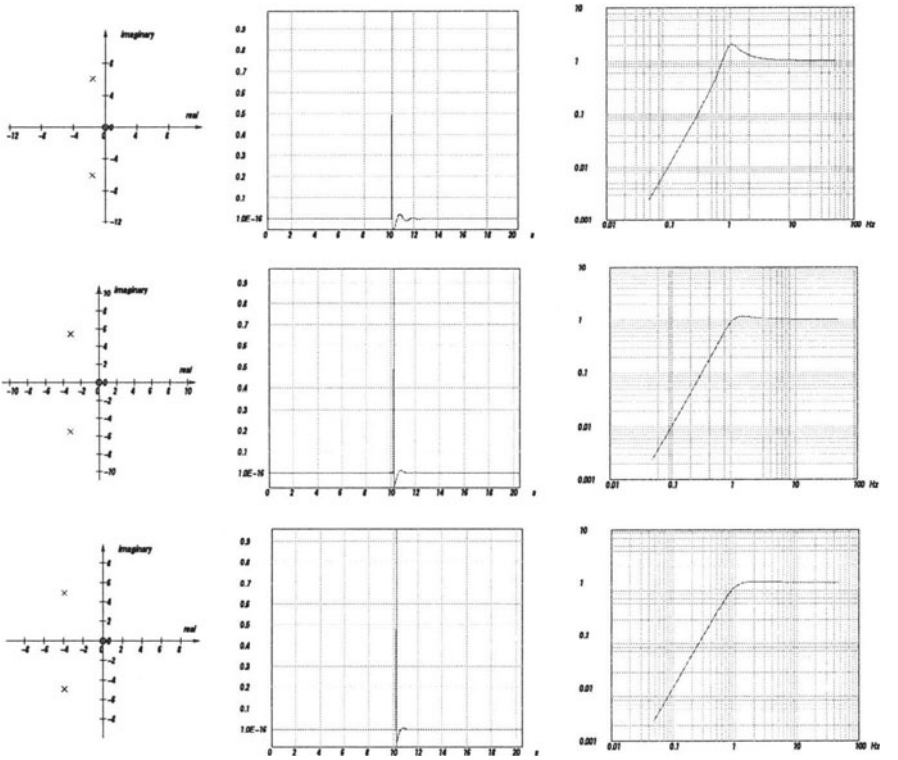


**Fig. A 4.1** Pole/zero maps (left column), impulse response functions (center column) and amplitude part of the frequency response functions (right columns) for a displacement seismometer with damping factors 0.25, 0.5, and 0.62 (from top to bottom). The internal sampling frequency in DST was set to 100 Hz (*Setup* menu).

You might notice some small precursory oscillations in the impulse response functions in Fig. A 4.1. The origin of this effect is discussed in detail in chapter 10.2.4.

*Solution 4.5* The output signal must be differentiated. This, however, corresponds to multiplication of the transfer function with $s$, a first degree polynomial with a single zero at the origin. Hence, in order to simulate the output of an electrodynamic system, simply add a zero at the origin to the transfer function. For the damping factor h = 0.62, the resulting amplitude frequency response is shown in Fig. A 4.2.

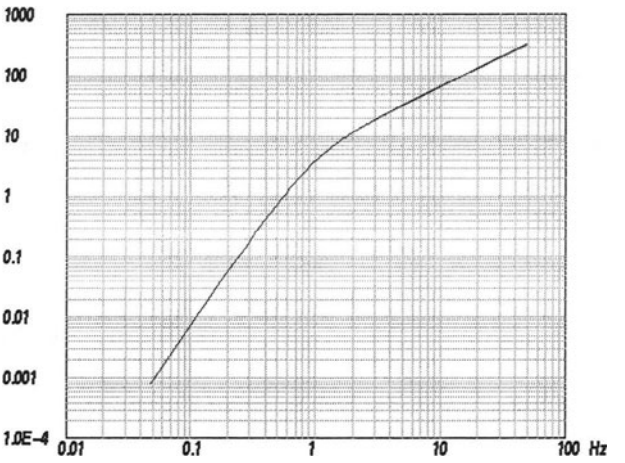



**Fig. A 4.2** Amplitude frequency response function for a 1 Hz velocity seismometer.

## Chapter 5

*Solution 5.1* First set the internal sampling frequency in DST to 1024 Hz and the window length to 2048 points (*Modify* option of the *Setup* menu). Next, generate the signals to be discretized using the *Sine/Cosine* option of the *Test Signals* menu. For the amplitude *A* and the phase *PHI* of the signal keep the value of 1 and 0, respectively. Start out with a signal frequency of 1 Hz and discretize and reconstruct this signal using a discretization frequency of 10 Hz. The result should agree with Fig. 5.2 - Fig. 5.4. The dominant frequency of the output signal can be easily obtained by measuring the dominant signal period on the DST screen in seconds and taking the reciprocal value. For the 1 Hz input signal the output signal frequency is 1 Hz also. Read the maximum signal amplitude of the output signal from the screen and enter input frequency, output frequency and maximum output amplitude into Table A 5.1.

**Table A 5.1** Corresponding input frequencies, output frequencies and output amplitudes for a discretization frequency of 10 Hz.

| Signal frequency "input signal" [Hz] | Signal frequency "output signal" [Hz] | Signal amplitude (maximum) "output signal" |
|---|---|---|
| 1 | 1 | $\approx 1$ |
| 2 | | |
| 3 | | |
| 4 | | |
| 5 | | |
| 6 | | |
| 7 | | |
| 8 | | |
| 9 | | |
| 10 | | |
| 11 | | |
| 12 | | |
| 13 | | |
| 14 | | |
| 15 | | |
| 16 | | |
| 17 | | |
| 18 | | |
| 19 | | |
| 20 | | |

Next, perform the same procedure with input signal frequencies from 2 - 20 Hz in steps of 1 Hz and fill in all entries in Table A 5.1.

For the interpretation imagine the "frequency band" as a band folded at 1/2 of the sampling frequency (folding frequency) and connect each input frequency with the corresponding output (alias-) frequency. As can be seen in Fig. A 5.1 the output frequencies can easily be obtained graphically by vertical projection from a particular input frequency on the "frequency band" down to the region between zero and the folding frequency.

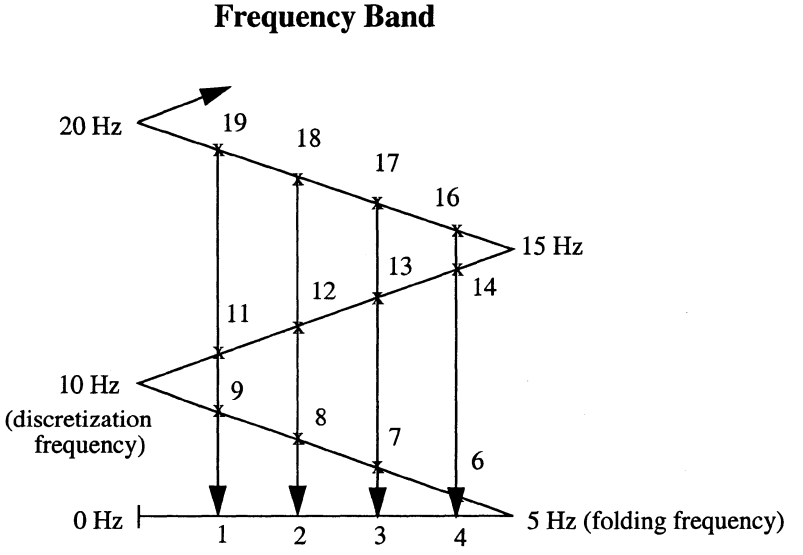



**Fig. A 5.1** Graphical illustration of the alias effect.

*Solution 5.2* For a discretization frequency of 10 Hz, the reconstruction breaks down for frequencies around 5 Hz. For exactly 5 Hz, it strongly depends on the value of the phase angle of the continuous signal, if the discretization results in non-zero values. For a phase angle of exactly zero, the signal is discretized exactly at the zero crossings. Hence it cannot be distinguished from a zero signal. Therefore, for a unique reconstruction the input signals have to be below 1/2 of 10 Hz (= 5 Hz).

*Solution 5.3* The aliasing frequency corresponding to 18.5 Hz and a sampling frequency of 10 Hz would be 1.5 Hz (Fig. A 5.2).

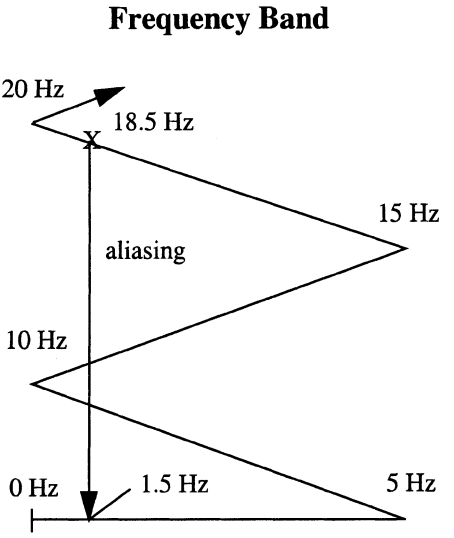



**Fig. A 5.2** Graphical determination of the alias frequency for a signal frequency of 18.5 Hz and a discretization frequency of 10 Hz..

## Chapter 6

*Solution 6.1* The Wood-Anderson magnitude is defined as $M_{WA} = \log_{10}(A) - \log_{10}(A_0)$ with $A$ the amplitude [in mm] measured on a Wood-Anderson displacement instrument and $-\log_{10}(A_0)$ the distance correction which is exactly 3 for 100 km. The theoretical trace amplitude for a $M_{WA} = 0$ earthquake on this instrument will be $A_{M_{WA}=0} = 10^{0-3} = 0.001\,\text{mm}$. For the magnitude 6 earthquake this value will be $A_{M_{WA}=0} = 10^{6-3} = 1000\,\text{mm}$. Hence the dynamic range needed is at least $20 \cdot \log(10^6) = 120$ dB (20 bits).

*Solution 6.2* The maximum number we can express with 16 bit is $2^{16} - 1$. For a gain ranging ADC with 8 bit for the gain and 8 bit for the mantissa we can express the maximum number $(2^8 - 1) \cdot 2^{(2^8 - 1)}$. Therefore, with the particular gain ranging setting we

increase the dynamic range by the factor

$$\frac{(2^8-1)\cdot 2^{(2^8-1)}}{(2^{16}-1)} = 2.25276\cdot 10^{74} .$$

Expressed in the decibel scale, this would yield an improvement of 1487 dB. Notice that this is a fictitious improvement since on the other hand, the loss of 8 bit of resolution would be unacceptable in most practical applications.

*Solution 6.3* First, set the internal sampling frequency in DST to 100 Hz and the window length to 512 points (*Modify* option of the *Setup* menu). Next, generate the signal to be discretized using the *Spike(s)* option of the *Test Signals* menu. For the spike position enter a value of 256 to center the spike in the middle of the window. For the amplitude enter a value of 100. Subsequently integrate this signal (*Utilitities -> Integration*) to produce a centered step funtion with a step amplitude of 1. Next, apply delta modulation (*Utilities-> Delta Modulation*) using a quantum value (LSB ) of 0.01 to obtain the following picture

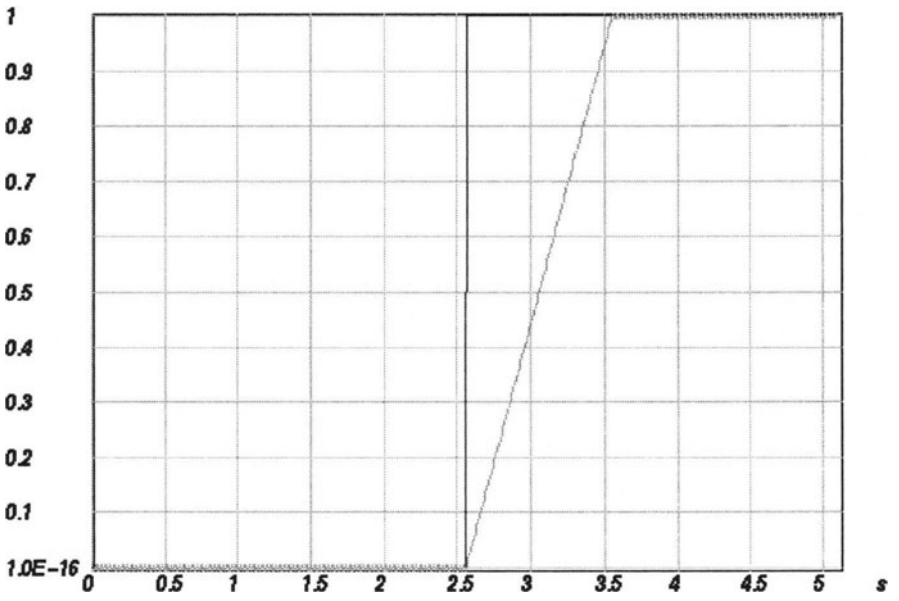



**Fig. A 6.1** Demonstration of slope overload and granular noise during delta modulation.

For the chosen parameters, the predicted input signal is unable to follow the rapid rise of the actual input signal. This effect is called *slope overload*. For the flat portions of the input signal, the predicted input signal oscillates around the actual input signal causing *granular noise*.

*Solution 6.4* In Fig. A 6.2 the performance of plain delta modulation is compared to delta modulation of the integrated signal omitting the integration in the demodulation part.

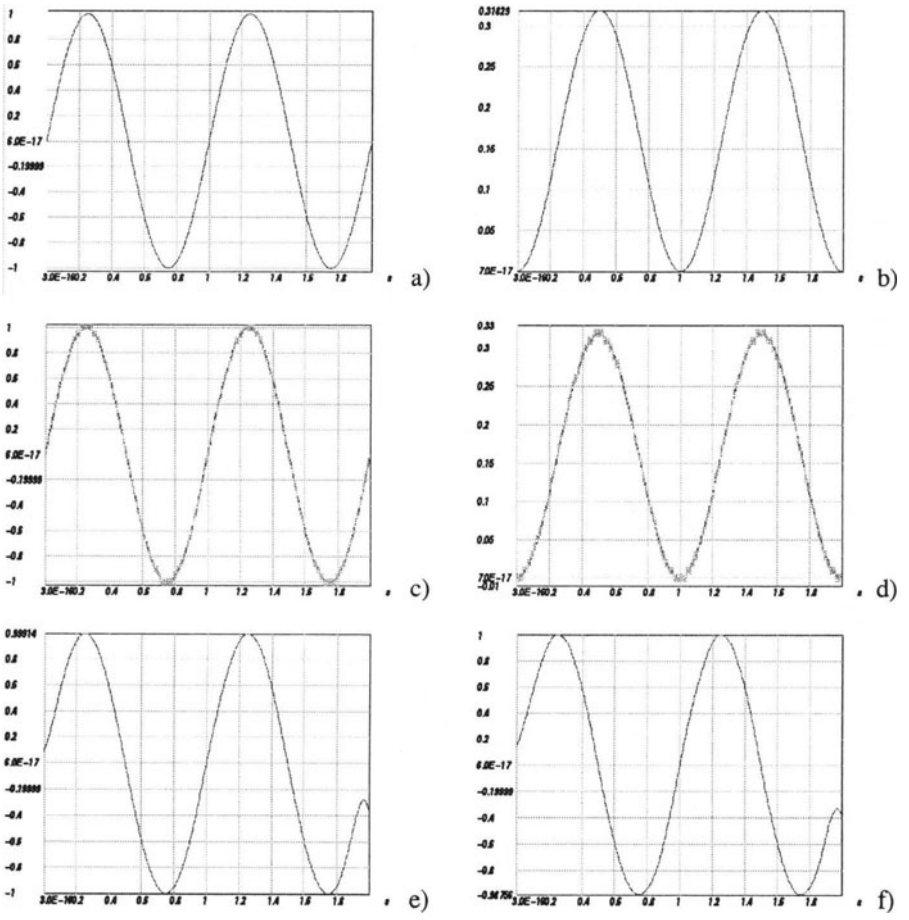



**Fig. A 6.2** Delta modulation and demodulation of a sinusoidal input signal with integration performed during demodulation (a, c, e) or before modulation (b, d, f).

The left panel in Fig. A 6.2 shows the sinusoidal input signal as defined in Problem 6.4 (a), the comparison of the actual input signal and predicted input signal (c), and the demodulated signal (e). The right hand panel shows in frame (b) the integrated sinusoid of Fig. A 6.2a), in frame (d) the comparison of the corresponding actual input signal and predicted input signal, and finally in frame (f) the low pass filtered modulator output signal. In the latter case the LSB value for the modulation of the integrated signal was reduced to 0.01 to account for the smaller maximum peak to peak amplitude. For the lowpass filter, in both cases a Butterworth filter with a corner frequency of 4 Hz and 1 section was chosen (*Utilities -> Butterworth Filter -> Lowpass*). As can be seen in Fig. A 6.2 the integration can be done either before modulation or before final lowpass filtering without changing the result. This could have been guessed already from the fact that integration is a linear operation.

*Solution 6.5* First set the internal sampling frequency in DST to 5120 Hz and the window length to 10240 points (*Modify* option of the *Setup* menu). Next, generate the 1 Hz sinusoidal signal to be discretized using the *Sine/Cosine* option of the *Test Signals* menu. For the amplitude *A* and the phase *PHI* of the signal take a value of 1 and 0, respectively. Use the *Sigma Delta Modulation* option from the *Utilities menu* of DST using a LSB of e. g. 1.2. The resulting trace will completely fill the screen. The final sampling frequency is generally reached in several stages. Here, use the *Decimation Filtering* option from the *Utilities* menu with a decimation ratio of 16 (resampling turned on) for the first stage. For the second stage use a decimation ratio of 4. Original sinusoid and resulting resampled trace are shown in Fig. A 6.3.

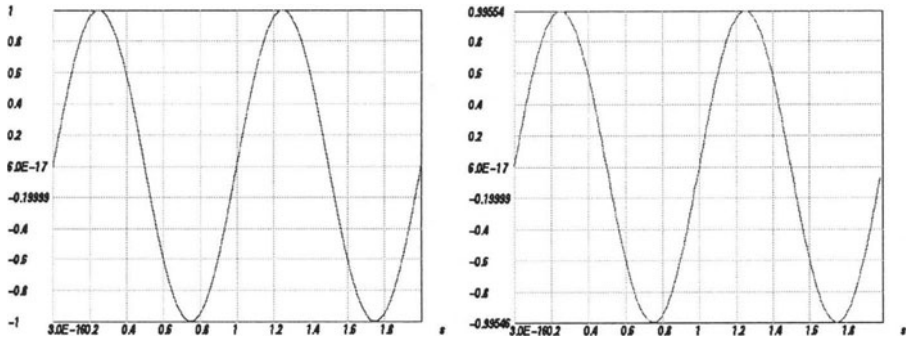



**Fig. A 6.3** Sigma-delta modulation simulation using DST. The left panel shows the input trace while the right panel displays the resulting resampled trace after sigma-delat modulation and two stages of decimation filtering.

## Chapter 7

*Solution 7.1* Since the signal is assumed to be zero outside the time window $0 \le t < NT$, the Fourier transform becomes $X(j\omega) = \int_0^{NT} x(t)e^{-j\omega t}\, dt$. If we approximate the integral by a finite sum using the rectangular rule we obtain:

$$\int_0^{NT} x(t)e^{-j\omega t}\, dt \approx T \cdot \sum_{n=0}^{N-1} x[nT]e^{-j\omega nT} \tag{7.1-1}$$

For $\omega = \omega_k = k \cdot \frac{2\pi}{TN}$ the right hand side becomes

$$\begin{aligned} T \cdot \sum_{n=0}^{N-1} x[nT]e^{-j\left(k \cdot \frac{2\pi}{TN}\right)nT} &= T \cdot \sum_{n=0}^{N-1} x[n]e^{-j\left(k \cdot \frac{2\pi}{N}\right)n} \\ &= T \cdot \tilde{X}[k] \end{aligned} \tag{7.1-2}$$

$\tilde{X}[k]$ is the DFT. Therefore,

$$X(j\omega)\big|_{\omega = \omega_k} = \int_0^{NT} x(t)e^{-j\omega_k t}\, dt \approx T \cdot \tilde{X}[k] \tag{7.1-3}$$

*Solution 7.2* Since the inverse Fourier transform evaluates the inverse z-transform on the unit circle, this means that the banded convergence region is chosen. Hence, it will correspond to a two-sided impulse response.

*Solution 7.3* a) positive time shifts $n_0$ will cause a multiplication of the z-transform with a term $z^{-n_0}$ which has a pole of order $n_0$ a at the origin ($1/0^{n_0}$). b) negative time shifts $n_0$ will cause a multiplication of the z-transform with a term $z^{n_0}$ which has a pole at infinity. Hence, depending on the sign of the shift, time shifts add poles at either the origin or at infinity of the z-plane.

*Solution 7.4* Let's assume we have two sequences with 1024 samples and a spike at position 1024 in each of them. If we convolve the two sequences in the time domain we expect the result to be a spike at position 2047. Hence, we have to pad up to this sample. If we have two sequences of different length, we have to pad both sequences up to one point less than twice the length of the longer sequence.

## Chapter 8

*Solution 8.1* In order to see how roots on the unit circle affect the properties of the impulse response of FIR filters, we first consider an impulse response function of length $M + 1 = 3$. Its z-transform in the notation of (8.18) corresponds to a simple quadratic

$q + pz + z^2$ with the two roots $z_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q} = re^{\pm j\Phi}$ for which the following two properties hold: $z_1 + z_2 = -p$ and $z_1 \cdot z_2 = q$ .

Since the wavelet is defined by the triplet ($q$, $p$, 1), we need to determine $q$ and $p$. For complex conjugate roots $re^{\pm j\phi}$ it follows that $q = z_1 \cdot z_2 = re^{j\phi} \cdot re^{-j\phi} = r^2$. Using Euler's formulas, we deduce that $-p = z_1 + z_2 = re^{j\phi} + re^{-j\phi} = 2r\cos\Phi$. Hence, the discrete finite impulse response for a system with the two complex conjugate roots $re^{\pm j\phi}$ will consist of the triplet

$$(q, p, 1) = (r^2, -2r\cos\phi, 1) \tag{8.1-1}$$

If the roots are on the unit circle ( $r = 1$ ), the result will be a *symmetric impulse response*. This will also be the case for higher order systems which only contain zeros on the unit circle since the convolution of symmetric impulse response functions remains symmetric. Because of this symmetry, one could think that in the context of removing all noncausal effects roots on the unit circle can be corrected for. However, this is not the case since a system with only roots on the unit circle has only a single waveform representation for the given amplitude spectrum. To see this we again consider an arbitrary second order system $A_1(z)$ with the two complex conjugate roots $z_{1,2} = r_1 e^{\pm j\phi_1}$ . Replacing the poles or zeros of a transfer function by their complex conjugate reciprocals does not change the 'amplitude spectrum' of the signal except for a constant scaling factor as we have already seen before. Therefore, the system $A_2(z) = A_1(1/z)$ with

the roots $z_{3,4} = 1/r_1 \cdot e^{\mp j\phi_1}$ has the same amplitude spectrum as $A_1(z)$ . In the general case in which the roots are assumed to be located away from the unit circle ( $r_1 \neq 1, r_2 \neq 1$ ) the waveforms of $A_1(z)$ and $A_2(z)$ will be different (cf. (8.1-1)). They are obtained by setting $r$ and $\phi$ in (8.1-1) to $r_1$ and $\phi_1$ and to $1/r_1$ and $-\phi_1$ for the first and the second representation, respectively. These are the only two waveform representations for the amplitude spectrum of $A_1(z)$ . One could naively argue that replacing only a single root by its reciprocal would not affect the amplitude spectrum as well. However, the roots of any real sequence of a second order system must always occur as a pair of complex conjugate roots, which would be violated in this case.

For higher order systems the number of wavelets with the same amplitude spectrum increases but this is irrelevant in the context of the present argument. Now we consider a second order system in which the roots are located directly on the unit circle. In this case, because $r_1 = 1$ the root pairs of $A_1(z)$ and $A_2(z)$ become identical

$z_{1,2} = z_{3,4} = e^{\pm j\Phi}$ . Consequently, the system has only a single waveform representation. The same argument can be made for higher order systems.

In the context of removing the acausal FIR filter response we can therefore ignore the contribution of roots on the unit circle (UC roots) because the corresponding waveform component will not be changed by the correction filter process. In effect, this is equivalent to treating the roots on the unit circle as belonging to the minimum phase component.

*Solution 8.2* Load the data trace example20HZ into DST using the *File -> Open File with Data Trace* option from the main menu. The file is found in subdirectory *FIR20HZ*. Next, use option *Interpolate* in the *Utilities* menu with an interpolation factor of 2 to obtain an interpolated trace sampled at 40 Hz. Subsequently use option *FIR2CAUS* in the *Utilities* menu with the correction filter file quant40Hz.prt. If you choose 0 for the number of zeros to pre/append/ to the filtered trace and select correction for linear phase, the resulting "corrected" trace is shown in Fig. A 8.1. Option Ok directly results in the corrected trace. If you choose *Apply* for the *FIR2CAUS* option, you will see both the uncorrected and the corrected trace displayed together. Using the left mouse button, you can zoom in on the P-wave onset as shown in Fig. A 8.1.

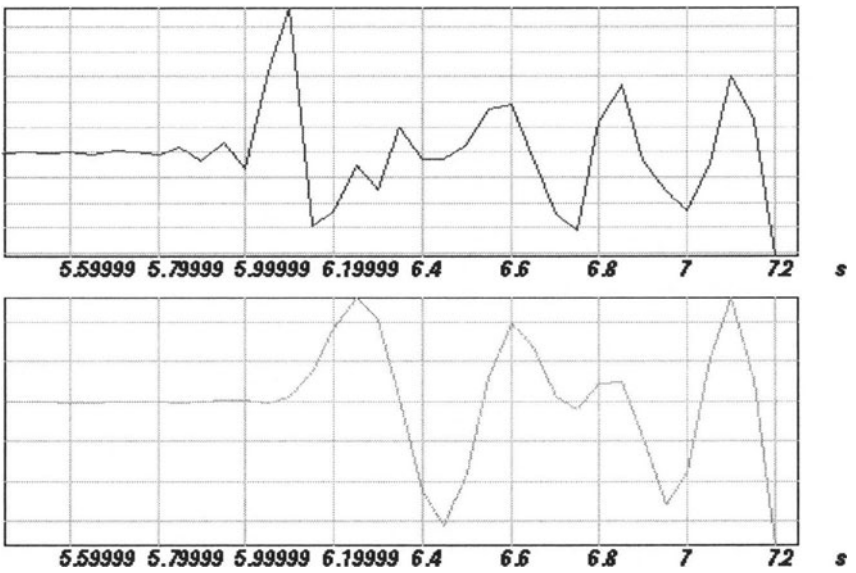



**Fig. A 8.1** Removing the maximum phase component of the FIR filter response from the example data trace example20HZ. The top trace shows the P- wave onset of the uncorrected trace while the bottom trace shows the corrected signal.

## Chapter 9

*Solution 9.1* In order to comply with the sampling theorem, digital recording systems must have an anti-alias filter for which the transfer function has to have zeros at the Nyquist frequency. Therefore, in addition to the irregularity at zero frequency discussed in the text, an irregularity problem will show up at the Nyquist frequency.

*Solution 9.2* Since a minimum phase system $T_{act}^{min}(z)$ has all its poles and zeros inside of the unit circle, this is true also for the inverse system $1/T_{act}^{min}(z)$. The poles of $T_{act}^{min}(z)$ have become zeros in $1/T_{act}^{min}(z)$ and vice versa. Hence, the inverse of a minimum phase system is always minimum phase as well. Since minimum phase systems are always stable (all poles are inside the unit circle since all singularities are inside the unit circle), **minimum phase systems always have stable causal inverse systems.**

$T_{act}^{max}(z)$ on the other hand as a maximum phase system has all its zeros outside the unit circle. These will become poles in $1/T_{act}^{max}(z)$ and as a consequence $1/T_{act}^{max}(z)$ will not have a stable causal impulse response.

*Solution 9.3* The continuous transfer function of the displacement seismometer is

$$T(s) = \frac{s^2}{s^2 + 2h_1\omega_{01}s + \omega_{01}^2} . \tag{9.3-1}$$

With frequency warping this becomes

$$T(s) = \frac{s^2}{s^2 + 2h_1\omega'_{01}s + \omega'^2_{01}} . \tag{9.3-2}$$

Replacing $s$ by $(2/T)\cdot(1-z^{-1})/(1+z^{-1})$ and $\omega'_{01}$ by $(2/T)\cdot\tan(\omega_{01}T/2)$ we obtain

$$T_{sim}(z) = \frac{\left(\frac{1-z^{-1}}{1+z^{-1}}\right)^2}{\left(\frac{1-z^{-1}}{1+z^{-1}}\right)^2 + 2h_1\tan\left(\omega_{01}\frac{T}{2}\right)\left(\frac{1-z^{-1}}{1+z^{-1}}\right) + \tan^2\left(\omega_{01}\frac{T}{2}\right)} \tag{9.3-3}$$

Multiplying numerator and denominator by $(1+z^{-1})^2$ yields

$$T_{sim}(z) = \frac{(1-z^{-1})^2}{(1-z^{-1})^2 + 2h_1\tan\left(\omega_{01}\frac{T}{2}\right)(1-z^{-1})(1+z^{-1}) + \tan\left(\omega_{01}\frac{T}{2}\right)^2(1+z^{-1})^2} \tag{9.3-4}$$

The numerator polynomial is simply $1-2z^{-1}+z^{-2}$, hence we get $b_0 = 1$, $b_1 = -2$, and $b_2 = 1$. We obtain the denominator polynomial by expanding the denominator in powers of $z^{-1}$ which becomes

$$
\begin{aligned}
&1 - 2z^{-1} + z^{-2} + 2h_1 \tan(\omega_{01}T/2) - 2h_1 \tan(\omega_{01}T/2)z^{-2} \\
&+ \tan^2(\omega_{01}T/2) + 2\tan^2(\omega_{01}T/2)z^{-1} + \tan^2(\omega_{01}T/2)z^{-2} \\
&= \\
&1 + 2h_1 \tan(\omega_{01}T/2) + \tan^2(\omega_{01}T/2) \\
&+ (-2 + 2\tan^2(\omega_{01}T/2)) \cdot z^{-1} \\
&+ (1 - 2h_1 \tan((\omega_{01}T/2) + \tan^2(\omega_{01}T/2))) \cdot z^{-2}
\end{aligned}
\tag{9.3-5}
$$

Hence we obtain

$$
\begin{aligned}
a_0 &= 1 + 2h_1 \tan(\omega_{01}T/2) + \tan^2(\omega_{01}T/2) \\
a_1 &= -2 + 2\tan^2(\omega_{01}T/2) \\
a_2 &= 1 - 2h_1 \tan((\omega_{01}T/2) + \tan^2(\omega_{01}T/2))
\end{aligned}
\tag{9.3-6}
$$

For $\omega_{01} = 2\pi \cdot 0.008333[Hz] = 0.052358$, $h_1 = 0.707$, and $T = 0.05[s]$ we finally get

$$
\begin{aligned}
a_0 &= 1.00185 \\
a_1 &= -2.0 \\
a_2 &= 0.998151
\end{aligned}
\tag{9.3-7}
$$

The performance of the filter is demonstrated in Fig. A 9.1.

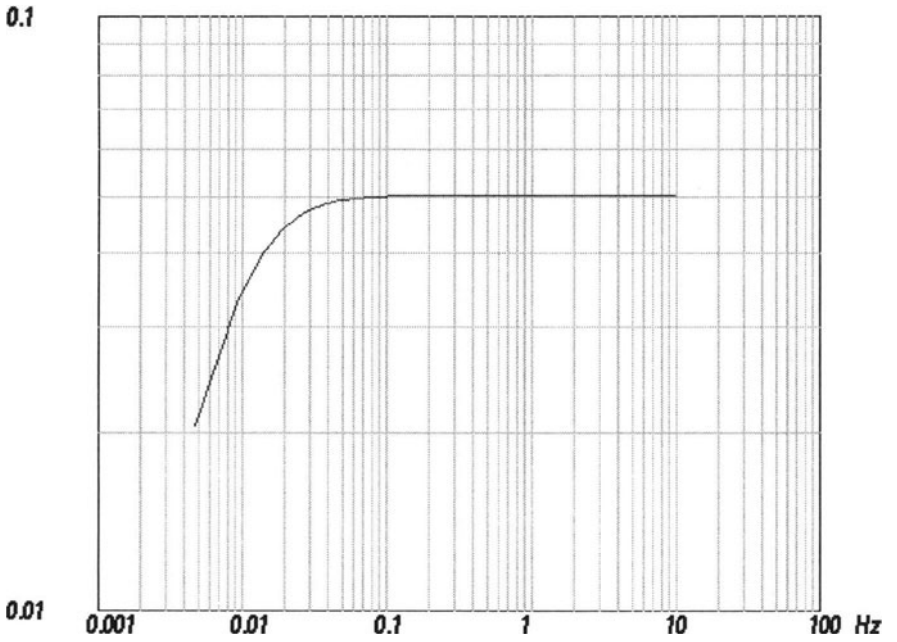



**Fig. A 9.1** Modulus of the frequency response function of a hypothetical seismometer with $h = 0.707$ and $\omega = 0.052358$.

Fig. A 9.1 was created within DST evaluating the filter output signal $y[i]$ from the filter input signal $x[i]$ and the AR coefficients $a[k]$ for $k = 0 - N$ and $b[l]$ for $l = 0 - M$ using following difference equation

$$y[i] = \sum_{k=1}^{N} -\left(\frac{a[k]}{a[0]} \cdot y[i-k]\right) + \sum_{l=0}^{M} \frac{b[l]}{a[0]} \cdot x[i-l] \tag{9.3-8}$$

First, the internal sampling frequency and the window length were set to 20 Hz and 4096 points, respectively. Next, a spike test signal was created at position 100 with an amplitude of 1 using the *Test Signals -> Spike(s)* option of DST. The resulting trace was filtered using the *Utilities -> Difference Equation* option in DST with autoregressive coefficients 1.00185, -2.0, 0.998151 and moving average coefficients 1, -2, 1. The gain factor was left at 1.0 and the *Filtering inverse option* was left off. Finally, Fig. A 9.1 was obtained using the *Utilities -> Spectrum -> Amplitude -> log-log* option of DST.

*Solution 9.4* The continuous transfer function of the inverse filter is

$$T^{-1}(s) = \frac{s^2 + 2h_1\omega_{01}s + \omega_{01}^2}{s^2} \tag{9.4-1}$$

With the frequency warping this becomes

$$T^{-1}(s) = \frac{s^2 + 2h_1\omega'_{01}s + \omega'^2_{01}}{s^2} \tag{9.4-2}$$

Replacing $s$ by $(2/T)\cdot(1-z^{-1})/(1+z^{-1})$ and $\omega'_{01}$ by $(2/T)\cdot\tan(\omega_{01}T/2)$ we obtain

$$T_{sim}(z) = \frac{\left(\frac{1-z^{-1}}{1+z^{-1}}\right)^2 + 2h_1\tan(\omega_{01}T/2)\left(\frac{1-z^{-1}}{1+z^{-1}}\right) + \tan^2(\omega_{01}T/2)}{\left(\frac{1-z^{-1}}{1+z^{-1}}\right)^2} \tag{9.4-3}$$

Multiplying numerator and denominator by $(1+z^{-1})^2$ yields

$$T_{sim}(z) = \frac{(1-z^{-1})^2 + 2h_1\tan(\omega_{01}T/2)(1-z^{-1})(1+z^{-1})}{(1-z^{-1})^2} + \frac{\tan^2(\omega_{01}T/2)(1+z^{-1})^2}{(1-z^{-1})^2} \tag{9.4-4}$$

The common denominator polynomial is simply $1-2z^{-1}+z^{-2}$, hence we get $a_0 = 1$, $a_1 = -2$, and $a_2 = 1$. We obtain the numerator polynomial by expanding the numerator in powers of $z^{-1}$ which becomes

$$
\begin{aligned}
&1 - 2z^{-1} + z^{-2} + 2h_1 \tan(\omega_{01}T/2) - 2h_1 \tan(\omega_{01}T/2)z^{-2} \\
&+ \tan^2(\omega_{01}T/2) + 2\tan^2(\omega_{01}T/2)z^{-1} + \tan^2(\omega_{01}T/2)z^{-2} \\
&= \\
&1 + 2h_1 \tan(\omega_{01}T/2) + \tan^2(\omega_{01}T/2) \\
&+ (-2 + 2\tan^2(\omega_{01}T/2)) \cdot z^{-1} \\
&+ (1 - 2h_1 \tan((\omega_{01}T/2) + \tan^2(\omega_{01}T/2))) \cdot z^{-2}
\end{aligned}
\tag{9.4-5}
$$

Hence we obtain

$$
\begin{aligned}
b_0 &= 1 + 2h_1 \tan(\omega_{01}T/2) + \tan^2(\omega_{01}T/2) \\
b_1 &= -2 + 2\tan^2(\omega_{01}T/2) \\
b_2 &= 1 - 2h_1 \tan((\omega_{01}T/2) + \tan^2(\omega_{01}T/2))
\end{aligned}
\tag{9.4-6}
$$

For $\omega_{01} = 2\pi \cdot 0.00833[Hz] = 0.052358$ , $h_1 = 0.707$ , and $T = 0.05[s]$ we finally get

$$
\begin{aligned}
b_0 &= 1.00185 \\
b_1 &= -2.0 \\
b_2 &= 0.998151
\end{aligned}
\tag{9.4-7}
$$

The performance of the filter is demonstrated in Fig. A 9.2.

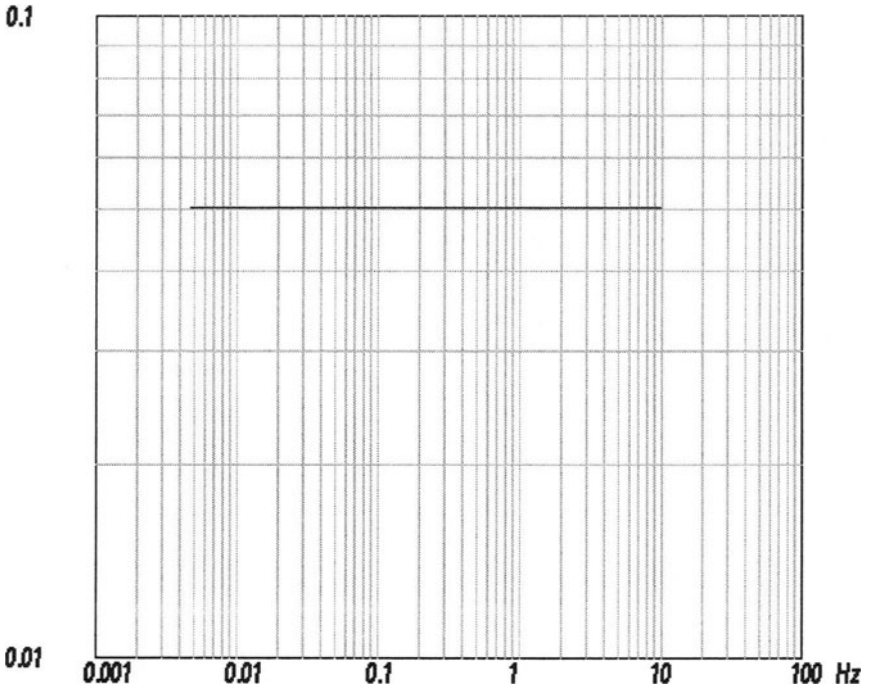



**Fig. A 9.2** Result of applying a deconvolution filter to the hypothetical seismometer response with $h = 0.707$ and $\omega = 0.052358$ shown in Fig. A 9.1. Displayed is the modulus of the resulting frequency response. The fact that the resulting frequency response is flat demonstrates the good performance of the deconvolution filter.

*Solution 9.5* In this case, the continuous transfer function of inverse filter is

$$T^{-1}(s) = \frac{s^2 + 2h_1\omega_{01}s + \omega_{01}^2}{s^3} \tag{9.5-1}$$

With the frequency warping this becomes

$$T^{-1}(s) = \frac{s^2 + 2h_1\omega'_{01}s + \omega'^2_{01}}{s^3} \tag{9.5-2}$$

Replacing $s$ by $(2/T)\cdot(1-z^{-1})/(1+z^{-1})$ and $\omega'_{01}$ by $(2/T)\cdot\tan(\omega_{01}T/2)$ we

obtain

$$T_{sim}(z) = \frac{T}{2} \cdot \frac{\left(\frac{1-z^{-1}}{1+z^{-1}}\right)^2 + 2h_1 \tan\left(\omega_{01}\frac{T}{2}\right)\left(\frac{1-z^{-1}}{1+z^{-1}}\right) + \tan^2\left(\omega_{01}\frac{T}{2}\right)}{\left(\frac{1-z^{-1}}{1+z^{-1}}\right)^3} \tag{9.5-3}$$

Multiplying numerator and denominator by $(1+z^{-1})^3$ yields

$$\begin{aligned} T_{sim}(z) = & \frac{(1-z^{-1})^2 \cdot (1+z^{-1})}{(1-z^{-1})^3} \cdot \frac{T}{2} \\ & + \frac{2h_1 \tan(\omega_{01}T/2)(1-z^{-1})(1+z^{-1})^2}{(1-z^{-1})^3} \cdot \frac{T}{2} \\ & + \frac{\tan(\omega_{01}T/2)^2(1+z^{-1})^3}{(1-z^{-1})^3} \cdot \frac{T}{2} \end{aligned} \tag{9.5-4}$$

The common denominator polynomial is simply $1-3z^{-1}+3z^{-2}-z^{-3}$, hence we get $a_0 = 1$, $a_1 = -3$, $a_2 = 3$ and $a_3 = -1$. We obtain the numerator polynomial by expanding the numerator in powers of $z^{-1}$ which becomes

$$
\begin{aligned}
&(T/2)\{1 - z^{-1} - z^{-2} + z^{-3} \\
&+ 2h_1 \tan\left(\omega_{01}\frac{T}{2}\right) + 2h_1 \tan\left(\omega_{01}\frac{T}{2}\right) z^{-1} - 2h_1 \tan\left(\omega_{01}\frac{T}{2}\right) z^{-2} - 2h_1 \tan\left(\omega_{01}\frac{T}{2}\right) z^{-3} \\
&+ \tan^2\left(\omega_{01}\frac{T}{2}\right) + 3\tan^2\left(\omega_{01}\frac{T}{2}\right) z^{-1} + 3\tan\left(\omega_{01}\frac{T}{2}\right)^2 z^{-2} + \tan^2\left(\omega_{01}\frac{T}{2}\right) z^{-3}\} \\
&= \\
&(T/2)\{1 + 2h_1 \tan\left(\omega_{01}\frac{T}{2}\right) + \tan^2\left(\omega_{01}\frac{T}{2}\right) \\
&+ \left(-1 + 2h_1 \tan\left(\omega_{01}\frac{T}{2}\right) + 3\tan^2\left(\omega_{01}\frac{T}{2}\right)\right) \cdot z^{-1} \\
&+ \left(-1 - 2h_1 \tan\left(\omega_{01}\frac{T}{2}\right) + 3\tan^2\left(\omega_{01}\frac{T}{2}\right)\right) \cdot z^{-2} \\
&+ \left(1 - 2h_1 \tan\left(\omega_{01}\frac{T}{2}\right) + \tan^2\left(\omega_{01}\frac{T}{2}\right)\right) \cdot z^{-3}\}
\end{aligned}
\tag{9.5-5}
$$

Hence we obtain

$$
\begin{aligned}
b_0 &= T/2 + Th_1 \tan\left(\omega_{01}\frac{T}{2}\right) + (T/2)\tan^2\left(\omega_{01}\frac{T}{2}\right) \\
b_1 &= -T/2 + Th_1 \tan\left(\omega_{01}\frac{T}{2}\right) + (1.5T)\tan^2\left(\omega_{01}\frac{T}{2}\right) \\
b_2 &= -T/2 - Th_1 \tan\left(\omega_{01}\frac{T}{2}\right) + (1.5T)\tan^2\left(\omega_{01}\frac{T}{2}\right) \\
b_3 &= (T/2) - Th_1 \tan\left(\omega_{01}\frac{T}{2}\right) + (T/2)\tan^2\left(\omega_{01}\frac{T}{2}\right)
\end{aligned}
\tag{9.5-6}
$$

For $\omega_{01} = 2\pi \cdot 0.00833[Hz] = 0.052358$ , $h_1 = 0.707$ , and $T = 0.05[s]$ we finally get

$$
\begin{aligned}
b_0 &= 0.0250463 \\
b_1 &= -0.0249536 \\
b_2 &= -0.0250461 \\
b_3 &= 0.0249538
\end{aligned}
\tag{9.5-7}
$$

The performance of the filter is demonstrated in Fig. A 9.3 with the P wave window of the vertical component record of the deep Bolivia earthquake of 94/06/09 recorded at station FUR of the German Regional Seismic Network. The data file FUR_940609.20HZ.Z is found in subdirectory FUR using the *File -> Open File with Data Trace* option from the main menu of DST. Here only the P wave onset is displayed.

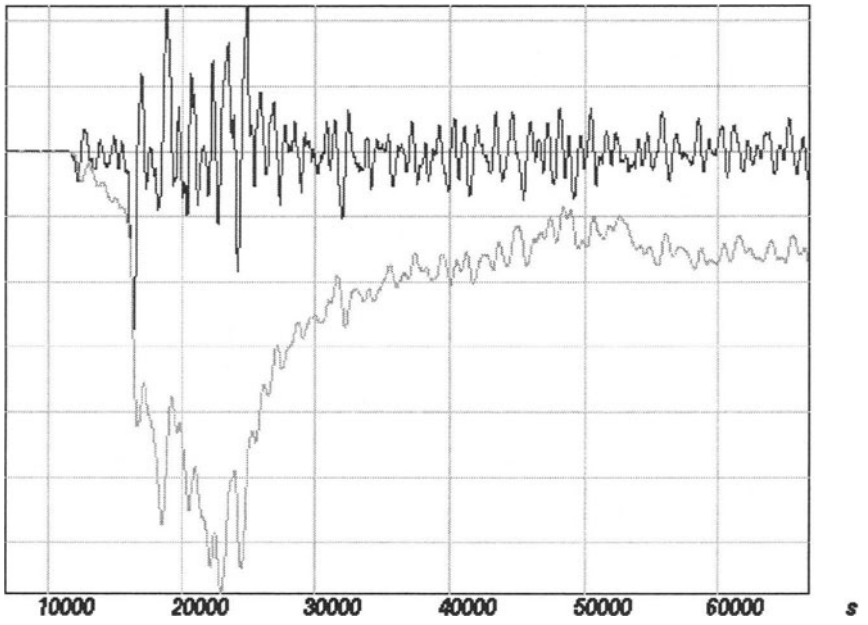



**Fig. A 9.3** P wave window of the vertical component record of the deep Bolivia earthquake of 94/06/09 recorded at station FUR of the German Regional Seismic Network. The top trace shows the BB 'velocity' record while the bottom trace shows the result of applying the deconvolution filter.

## Chapter 10

*Solution 10.1* These signals can be reproduced within DST by producing a spike test signal for an internal sampling frequency of 20 Hz and a trace length of 256 points and a spike of amplitude 1 at position 10. Trace 1 was produced by causal bandpass filtering the spike trace with a Butterworth bandpass consisting of 3 sections (6 poles) and corner frequencies at 0.4 and 1.5 Hz. Trace 2 corresponds to the spike trace filtered by a 6 pole Butterworth lowpass filter with a cut off frequency of 1.5 Hz. Trace 3 finally consists of a sinusoidal signal of 1 Hz signal frequency, amplitude of 1, and zero phase, again for 20 Hz sampling frequency and a trace length of 1024 points.

The frequency response function is defined by the following 9 poles:

-0.20762 0
-0.71256 1.03508
-0.71256 -1.03508
-7.18 +17.42
-7.18 -17.42
-17.42 +7.18
-17.42 -7.18
-37.04 0
-45.45 0

and 4 zeros at the origin. The corresponding GSE calibration file can be loaded into DST using the *File -> Open File With GSE Calibration* option from the main menu (file: pr10_1.cal). Looking at the amplitude frequency response function in Fig. A 10.1, we notice approximately three regions with different slope in the log-log plot. Outside of the central frequency band, roughly below 0.2 and above 2.5 Hz, the amplitude frequency response function decays rapidly, hence signals with frequencies outside this range are strongly attenuated while signals with frequencies between 0.2 and 2.5 Hz could be well recorded. With what we have learned about seismometer systems, we could expect that the frequencies of 0.2 and 2.5 Hz correspond to the eigenfrequency of the seismometer and the corner frequency of the anti-alias filter, respectively. This frequency band is commonly defined as the passband of the instrument. Within the passband, we see that the amplitude frequency response function is proportional to $\omega$ so from the correspondence between multiplication by $j\omega$ in the frequency domain and differentiation in the time domain we could expect that the frequency response function is that of an electrodynamic system. In this case, all signals falling completely within the passband would be expected to be differentiated. By comparing the spectra in Fig. 10.4 and the frequency response function in Fig. 10.5, we would expect this to happen for traces 1 and 3.

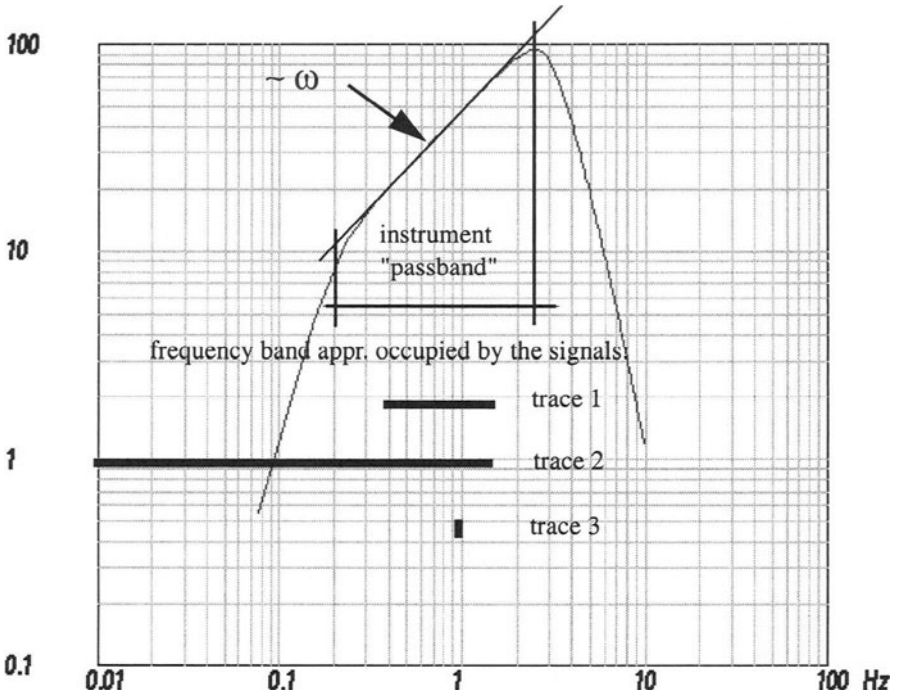



**Fig. A 10.1** Amplitude frequency response function for Problem 10.1.

The actual result for trace 1 is shown in Fig. A 10.2. The waveform of the filtered output signal is essentially identical to the scaled and differentiated input signals. Slight differences can be understood from the fact that the signal contains small amounts of energy outside the passband.

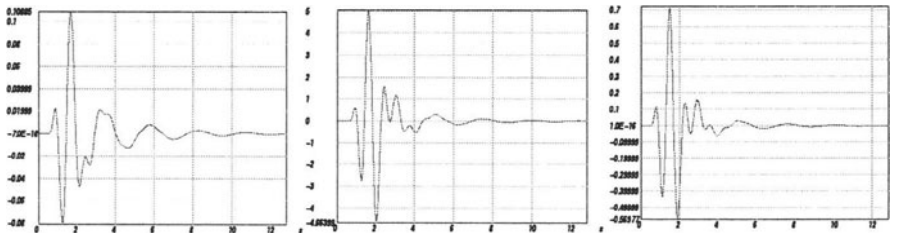


**Fig. A 10.2** The traces from left to right show the input signal trace 1 from Fig. 10.3, the output signal from filtering with the system displayed in Fig. A 10.1, and the differentiated input signal.

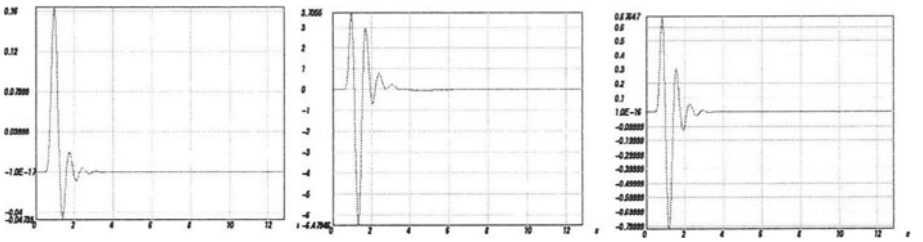


**Fig. A 10.3** The traces from left to right show the input signal trace 2 from Fig. 10.3, the output signal from filtering with the system displayed in Fig. A 10.1, and the differentiated input signal.

For trace 2 (Fig. A 10.3), the differences between the filtered trace and the differentiated trace are stronger since the input signal contains considerable amounts of energy outside the passband of the instrument (Fig. A 10.1).

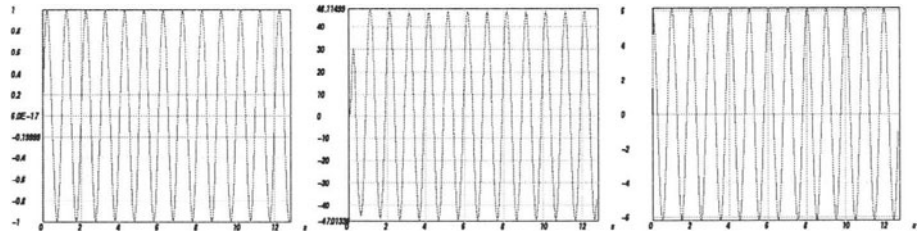


**Fig. A 10.4** The traces from left to right show the input signal trace 3 from Fig. 10.3, the output signal from filtering with the system displayed in Fig. A 10.1, and the differentiated input signal.

In Fig. A 10.4 it is shown how the fictitious instrument acts on a sinusoidal signal with 1.0 Hz signal frequency. What we can observe nicely is that the output signal is again a sinusoid with the same frequency. This is an important result which relates to the eigenvector property of the harmonic functions for LTI systems. Only amplitude and phase are changed. However, we can also see that there is a phase shift between trace 2 and 3. Hence, for the monochromatic signal with a signal frequency completely within the region of the frequency response function where its modulus is proportional to $\omega$, the output signal is not simply the differentiated signal which would show a phase shift to the input signal of $90^{\circ}$. If we calculate the phase shift for the given signal frequency of 1 Hz from the pole and zero distribution using the techniques learned earlier, we obtain approximately $37^{\circ}$ which clearly differs from $90^{\circ}$ which we would expect for pure differentiation.

How does this fit in with the fact that for the signal in Fig. A 10.4 the difference between the signal filtered with the instrument does not differ significantly from the differentiated

signal? The answer lies in the linear phase components of the frequency response function. Linear phase components do not cause any signal distortion but only a time shift of the output signal. With other words, we could also explain the phase shift seen for the monochromatic signal as being produced by a time shift and a differentiation. However, in order to test this hypothesis we would have to analyse the properties of the phase of the frequency response function in more detail. In this context, however, I only want to remind you that the circumstances under which it is justified to speak of a *displacement-* or a *velocity system* depend on a number of factors and that we should be careful not to use these terms sloppily.

*Solution 10.2* For a displacement frequency response function it would be counts/ displacement, for a velocity frequency response function it would be counts/velocity, and counts/acceleration for an acceleration frequency response function respectively.

For example for the displacement frequency response function, we have displacement as unit for the input signal, displacement per Hz for the spectrum, counts for the unit of the output signal and counts per Hz for the corresponding spectrum. Since the frequency response function is also defined as the Fourier spectrum of the output signal divided by the Fourier spectrum of the input signal, the unit is counts/displacement.

*Solution 10.3* The velocity impulse response function is the inverse Fourier transform of:

$$T_{vel}(j\omega) = \frac{Output(j\omega)}{Input_{vel}(j\omega)} \tag{10.3-1}$$

a) Displacement impulse response: What we need to know is the response of the system to an input signal given in displacement. By equivalence between differentiation in the time domain and multiplication with $j\omega$ in the frequency domain, the velocity input spectrum is related to the displacement input spectrum by

$$Input_{vel}(j\omega) = j\omega \cdot Input_{disp}(j\omega) \tag{10.3-2}$$

Hence

$$T_{vel}(j\omega) = \frac{Output(j\omega)}{j\omega \cdot Input_{disp}(j\omega)} \tag{10.3-3}$$

and replacing $Output(j\omega)/Input_{disp}(j\omega)$ by $T_{disp}(j\omega)$ we obtain

$$T_{disp}(j\omega) = \frac{Output(j\omega)}{Input_{disp}(j\omega)} = T_{vel}(j\omega) \cdot j\omega \tag{10.3-4}$$

Hence, the displacement frequency response function can be obtained from the velocity frequency response function by multiplication with $j\omega$. Again, by equivalence between multiplication with $j\omega$ in the frequency domain and differentiation in the time domain, the displacement impulse response function is obtained from the velocity impulse response function by differentiation.

b) Acceleration impulse response: By the same argument we obtain the acceleration frequency response function from the velocity response function by division by $j\omega$. Finally, by equivalence between division by $j\omega$ in the frequency domain and integration in the time domain, the acceleration impulse response function can be obtained from the velocity impulse response function by integration.

*Solution 10.4* We can directly use (10.7) with $\omega_0 = 2\pi \cdot f_0 = 0.0519657$, h = 0.718, and G = 1500 [V/m/s] to obtain the velocity transfer function

$$T_{vel}(s) = -1500[V/m/s]\frac{s^2}{s^2 + 0.07462s + 0.0027} \tag{10.4-1}$$

From the discussion of Problem 10.3 we know that we obtain the displacement response function by multiplication with $s$.

$$T_{disp}(s) = -1500[V/m]\frac{s^3}{s^2 + 0.07462s + 0.0027} \tag{10.4-2}$$

The unit of the scale factor in this case becomes [V/m]. We can also express the denominator polynomial by its roots (see equation (10.5)) which are given by equation (10.8) ( $s_{p(1,2)} = -(h \pm \sqrt{h^2 - 1}) \cdot \omega_0$ ). With the parameters given above we obtain

$s_{p(1.,2)} = -(0.718 \pm 0.696 \cdot j) \cdot 0.0519657 = -(0.03731 \pm 0.03617 \cdot j)$ and

$$T_{disp}(s) = \frac{-1500\left[\frac{V}{m}\right] \cdot s^3}{(s + 0.03731 + j \cdot 0.03617) \cdot (s + 0.03731 - j \cdot 0.03617)} \tag{10.4-3}$$

The corresponding frequency response function is obtained by evaluating $T(s)$ on the

imaginary axis ($s \rightarrow j\omega$).

$$T_{disp}(j\omega) = \frac{-1500\left[\frac{V}{m}\right].}{(j\omega)^3} \qquad (10.4\text{-}4)$$
$$(j\omega + 0.03731 + j \cdot 0.03617) \cdot (j\omega + 0.03731 - j \cdot 0.03617)$$

*Solution 10.5* With an LSB of 2.5 $\mu V$ we obtain the $1V = 0.4 \cdot 10^6$ counts. Hence $T_{disp}(j\omega)$ has to be multiplied by $0.4 \cdot 10^6$ to come out in counts and to be divided by $10^9$ to convert $m$ into $nm$.

$$T_{disp}(j\omega) = \frac{-0.6\left[\frac{cts}{nm}\right].}{(j\omega)^3} \qquad (10.5\text{-}1)$$
$$(j\omega + 0.03731 + j \cdot 0.03617) \cdot (j\omega + 0.03731 - j \cdot 0.03617)$$

The shape of the modulus of the displacement frequency response function for this system is shown in Fig. A 10.5.

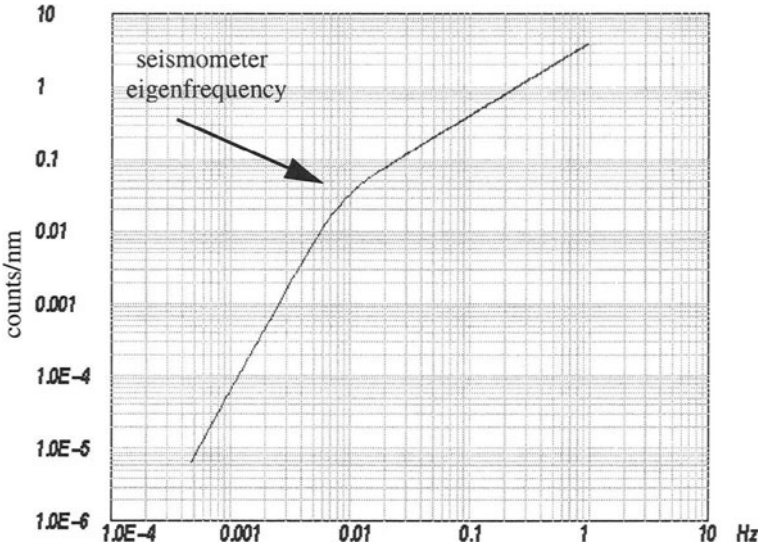



**Fig. A 10.5** Displacement amplitude frequency response function for Problem 10.5. Amplitude in counts/nm.

*Solution 10.6* As discussed in the context of Problem 10.3, the displacement frequency response function can be obtained from the velocity response function by multiplication with $j\omega$. Again, by equivalence between multiplication with $j\omega$ in the frequency domain and differentiation in the time domain, the displacement impulse response function is obtained from the velocity impulse response function by differentiation.

Here we want to go the other way, hence in order to obtain the velocity frequency response function, we have to either calculate the spectrum from the integrated displacement impulse response function or divide the displacement frequency response function by $j\omega$. Another way is by removing a zero from the GSE calibration file describing the transfer function. The result is shown in Fig. A 10.6.

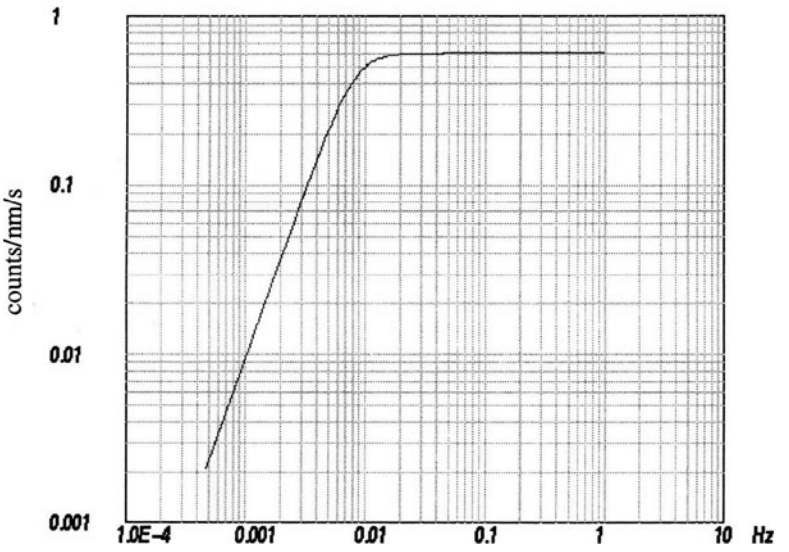



**Fig. A 10.6** Velocity amplitude frequency response function for Problem 10.6. Amplitudes in counts/nm/s.

Note that the plateau value is 0.6 counts/nm/s, which corresponds to the generator constant expressed in counts[1].

*Solution 10.7* The scale factor in equation (10.11)

$$T(j\omega) = T_{disp}(j\omega) = C_{disp} \cdot \frac{\prod_{k=1}^{L} (j\omega - s_{0k})}{\prod_{k=1}^{N} (j\omega - s_{pk})} = C_{disp} \cdot F_{disp}(j\omega) \qquad (10.7\text{-}1)$$

must be chosen such that $T_{disp}(j\omega_{cal}) = 1/g_d$.

Hence we obtain $1/g_d = C_{disp} \cdot |F_{disp}(j\omega_{cal})|$ which yields

1. The systems discussed in Problem 10.4 to Problem 10.6 describes the stations of the German Regional Seismic Network (GRSN) as of June 1994.

$$C_{disp} = 1/(g_d \cdot |F_{disp}(j\omega_{cal})|) \ . \qquad (10.7\text{-}2)$$

*Solution 10.8* In the context of the solution of Problem 10.7, we have seen that the scale factor $C_{disp}$ is calculated as $C_{disp} = 1/(g_d \cdot |F_{disp}(j\omega_{cal})|)$. Here, $g_d$ was defined as $g_d = A_i^{disp}/A_o$, with $A_i^{disp}$ the displacement amplitude of the input signal and $A_o$ being the amplitude of the output signal in counts. If we are measuring the velocity amplitude of the input signal as in Fig. A 10.7, we get $g_v = A_i^{vel}/A_o$ or in case of acceleration $g_a = A_i^{acc}/A_o$.

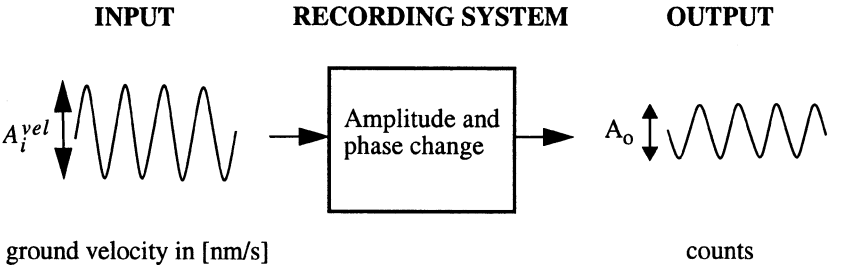



**Fig. A 10.7** Sketch of a shaking table experiment to measure the calibration gain $g_v$ at a given period.

However, for an harmonic signal at calibration frequency $\omega_{cal}$, we obtain $A_i^{acc} = \omega_{cal} \cdot A_i^{vel} = \omega_{cal}^2 \cdot A_i^{disp}$ and $A_i^{vel} = \omega_{cal} \cdot A_i^{disp}$. Therefore, we get $g_v = \dfrac{A_i^{vel}}{A_o} = \dfrac{\omega_{cal} \cdot A_i^{disp}}{A_o} = \omega_{cal} \cdot g_d$ and $g_d = \dfrac{g_v}{\omega_{cal}}$ for the relationship between $g_v$ and $g_d$. Likewise, $g_d = \dfrac{g_a}{\omega_{cal}^2}$. As a consequence, we obtain for the scale factor

$$C_{disp} = \frac{1}{g_d \cdot |F(j\omega_{cal})|} = \frac{\omega_{cal}}{g_v \cdot |F(j\omega_{cal})|} = \frac{\omega_{cal}^2}{g_a \cdot |F(j\omega_{cal})|} \qquad (10.8\text{-}1)$$

*Solution 10.9* Since the sensor is supposed to generate a voltage which is proportional to ground velocity for frequencies above 1 Hz, we have a system similar to that described in Solution 10.4 only with different parameters. For $G = 100$ V/m/s, $f_0 = 1.0$ Hz ($\omega_0 = 6.283$), and h = 0.7 the velocity transfer function is

$$T_{vel}(s) = -100\left[\frac{V}{m/s}\right]\frac{s^2}{s^2 + 8.7964s + 39.476} \qquad (10.9\text{-}1)$$

From the discussion of Problem 10.3 we know that we can calculate the displacement response function by multiplication with $s$

$$T_{disp}(s) = -100\left[\frac{V}{m}\right]\frac{s^3}{s^2 + 8.7964s + 39.476} \qquad (10.9\text{-}2)$$

The unit of the scale factor in this case becomes [V/m]. We can express the denominator polynomial by its roots (see equation (10.5)) which are given by equation (10.8) ( $s_{p(1,2)} = -(h \pm \sqrt{h^2 - 1}) \cdot \omega_0$ ). With the parameters given above we obtain

$$s_{p(1.,2)} = -(0.7 \pm j \cdot 0.71414) \cdot 6.283 = -(4.398 \pm j \cdot 4.487) \qquad (10.9\text{-}3)$$

As in Solution 10.4 we can use equation (10.5) with these 2 poles and three zeros at the origin and the fact that $\beta_L = 100$ [V/m] and $\alpha_N = 1$, hence $(\beta_L/\alpha_N) = 100$ [V/m] in (10.5). However, since the signal is amplified by 250 before being fed into the A/D converter, we must be aware that the effective factor $\beta_L$ is $100 \cdot 250 = 2.5 \cdot 10^4$. For the frequency response function (evaluating $T(s)$ for $s = j\omega$) relating output voltage and ground displacement we obtain

$$T_{disp}(j\omega) = \frac{-2.5 \cdot 10^4\left[\frac{V}{m}\right].}{\frac{(j\omega)^3}{(j\omega + 4.398 + j \cdot 4.487) \cdot (j\omega + 4.398 - j \cdot 4.487)}} \qquad (10.9\text{-}4)$$

With an LSB of 1 $\mu V$ we obtain $1V = 10^6$ counts. Hence $T_{disp}(j\omega)$ has to be multiplied by $10^6$ to come out in counts and to be divided by $10^9$ to convert $m$ into $nm$.

$$T_{disp}(j\omega) = \frac{-25 \cdot \left[\frac{counts}{nm}\right] \cdot (j\omega)^3}{(j\omega + 4.398 + j \cdot 4.487) \cdot (j\omega + 4.398 - j \cdot 4.487)} \tag{10.9-5}$$

Following the argument in Solution 10.6, we obtain the velocity frequency response function by dividing the displacement frequency response function by $j\omega$.

$$T_{vel}(j\omega) = \frac{-25 \cdot \left[\frac{counts}{nm}\right] \cdot (j\omega)^2}{(j\omega + 4.398 + j \cdot 4.487) \cdot (j\omega + 4.398 - j \cdot 4.487)} \tag{10.9-6}$$

In order to obtain $g_d = 1/|T_{disp}(j\omega_{cal})|$ and $g_v = 1/|T_{vel}(j\omega_{cal})|$, we have to evaluate the modulus of the displacement frequency response function and the velocity response function at the calibration frequencies. Calculating these values (using a pocket calculator) and $f_{cal} = \omega_{cal}/2\pi$ of 5 and 10 Hz, respectively, we obtain:

5 Hz: $g_d = 0.00127324$ [nm/counts] and $g_v = 0.03999$ [nm/s/counts]

10 Hz: $g_d = 0.00063652$ [nm/counts] and $g_v = 0.03999$ [nm/s/counts].

From Solution 10.8 we know that for harmonic input signals with signal frequency equal to the calibration frequency we obtain $g_d = A_i^{disp}/A_o$ and $g_v = A_i^{vel}/A_o$. Hence the frequency response function at the calibration frequency should be $A_o = A_i^{disp}/g_d$ and $A_o = A_i^{vel}/g_v$.

In Fig. A 10.8, the displacement frequency response function has been calculated using DST for the two poles at $4.398 \pm 4.487$, three zeros at the origin and a scale factor of 25 (internal sampling frequency 50 Hz, window length 4096).

So for the following argument $A_i = 1.0$. At 5 Hz, we expect $A_o = A_i^{disp}/g_d = 1.0/0.00127324 = 785$ [counts]. Likewise at 10 Hz we obtain $A_o = 1571$ [counts] which can be seen to be the case in Fig. A 10.8.

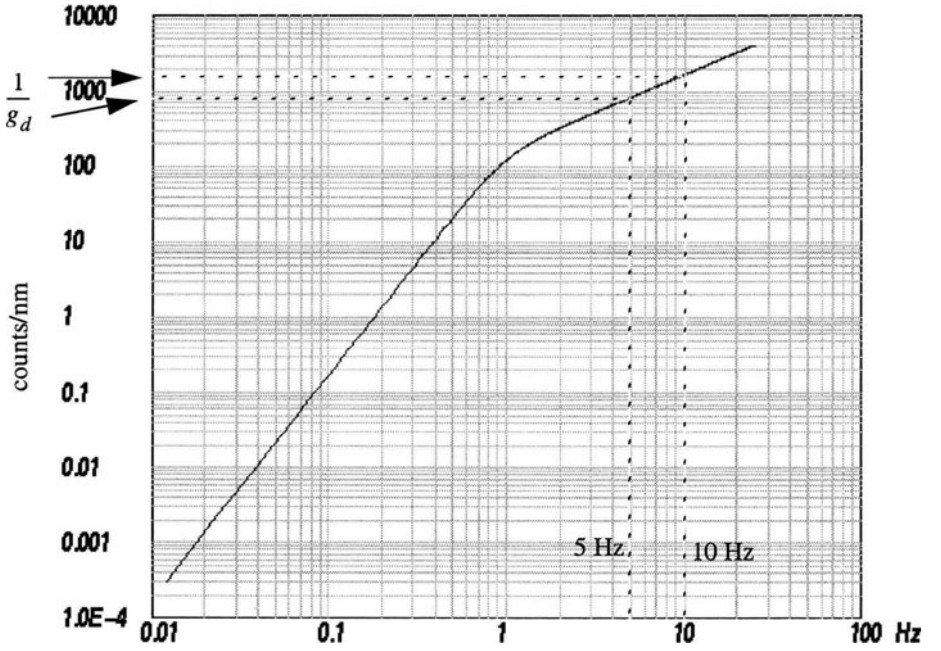



**Fig. A 10.8** Displacement amplitude frequency response function for Problem 10.9.

In Fig. A 10.9, the velocity frequency response function is displayed. It has been calculated the same way as the displacement frequency response function except for the fact that one of the zeros has been removed in order to account for the differences between displacement and velocity frequency response function.

From the same argument as for the displacement frequency response at 5 Hz, we expect $A_o = A_i^{vel}/g_v = 1.0/0.03999 = 25$ counts and the same value as for 10 Hz. We see in Fig. A 10.9 that both calibration frequencies are within the flat portion of the velocity frequency response function.

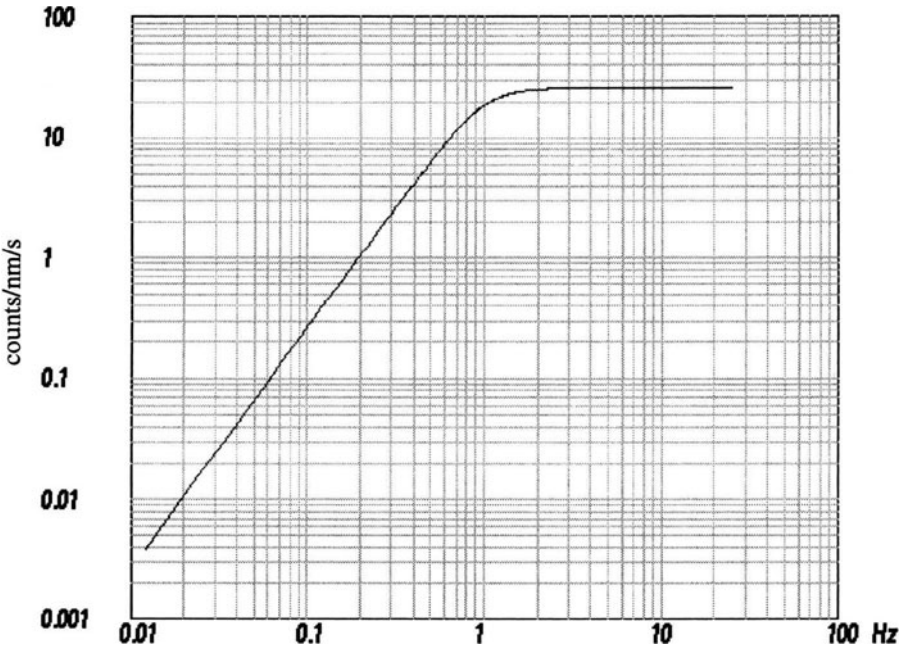



**Fig. A 10.9** Velocity amplitude frequency response function for Problem 10.9.

*Solution 10.10* Since the signal is a harmonic signal, all we have to do is determine the magnification in [counts/nm] for ground displacement at the signal frequency and divide the peak-to-peak amplitude in counts by that factor. Since we have 5 oscillations within 1 sec, the signal frequency is 5 Hz. We have already calculated the magnification at that frequency (Solution 10.9). For the peak-to-peak amplitude of 396 counts, we therefore obtain the peak-to-peak displacement amplitude as:

$$396/\left|T_{disp}(j\cdot 2\pi f)\right|_{f\,=\,5Hz} \;=\; 396\cdot g_d^{5Hz} \;=\; 396\cdot 0.00127324 \;=\; 0.5042nm \quad .$$

*Solution 10.11* Because the resulting filter is causal, there is no signal before the sample corresponding in time to the input 'spike'. As a consequence, the zero crossings of the harmonic components at the Nyquist frequency must remain exactly at sample intervals. Therefore, the corresponding phase delay at the Nyquist frequency must be an multiple of the sampling interval, which necessarily causes a distortion of the phase response.

*Solution 10.12* Because the slope of a sinusoid $A_0 \cdot \sin(\omega t)$ equals $\omega \cdot A_0 \cdot \cos(\omega t)$, $slope_{max} = \omega \cdot A_0$. and with $a_{max} = 2A_0$ the rise time for a sinusoid becomes

$$t_{r(\text{sinusoid})} = 2/\omega = 2/(2\pi \cdot f) = T/\pi \qquad (10.12\text{-}1)$$

with $T$ the period of the sinusoid. Hence, for a sinusoidal signal, the rise time and the duration as measured from the time difference between zero crossings are related by a constant factor of $\pi$.

*Solution 10.13* In Fig. A 10.10 the input spike, the causal and the acausal impulse responses of the 8 pole (4 section), 1 Hz Butterworth LP filter are displayed.

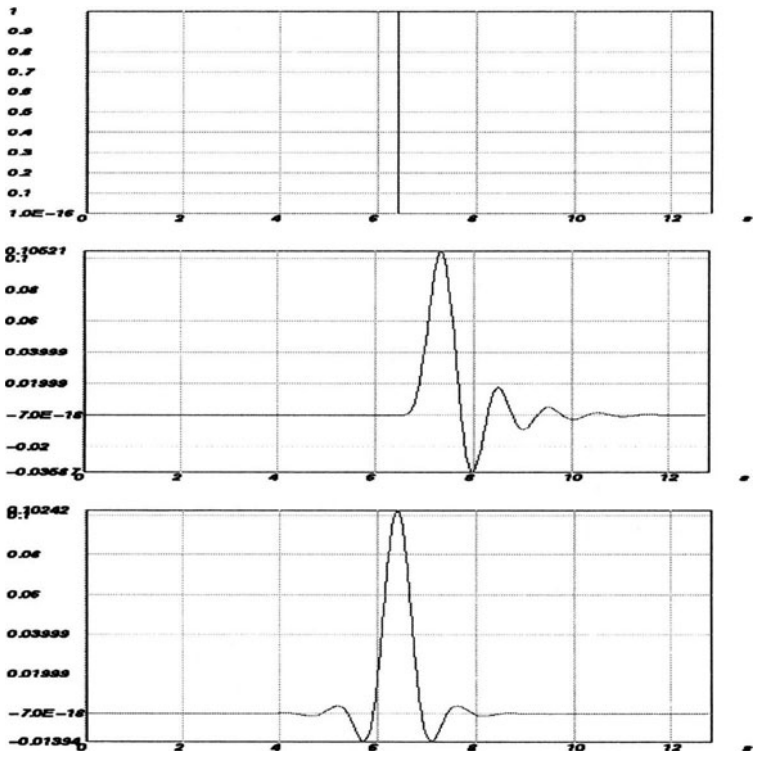


**Fig. A 10.10** Input spike, the causal and the acausal impulse response of the 8 pole (4 section) Butterworth filter (from top to bottom)

In contrast to the acausal filter, the causal filter shows a delay of both the maximum amplitude and the onset. For the acausal filter, the maximum amplitude is not delayed.

Several properties can be noted. For the causal filter, the onset shows the change in the onset order which slightly enlarges the pulse duration as measured from zero crossings with respect to the acausal filter. Remember that for the acausal filter the definition of an onset makes no sense. While the oscillations in the impulse response of the causal filter have larger amplitudes the frequencies are the same. Similar effects are visible for the step response function in Fig. A 10.11. Since the maximum values of the impulse response functions are very close, the maximum slopes of the step response functions and the rise times will be similar as well. Nevertheless, visually we note an apparent increase in the rise time for the causal filter, again caused by the change of the onset order.

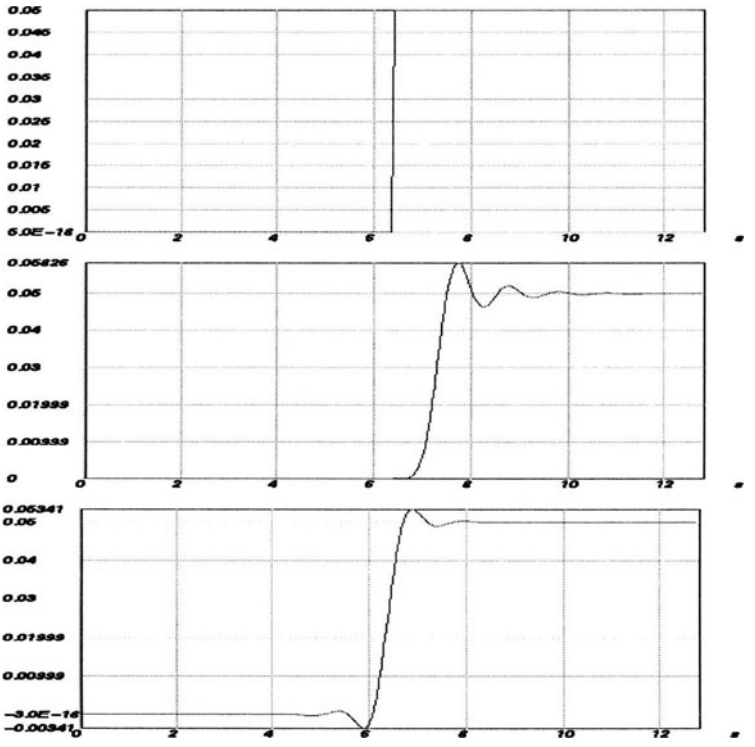



**Fig. A 10.11** Superimposed step response functions for the 8 pole, 1Hz Butterworth LP filter for the causal and acausal case.

# Index


**A**
ADC 73
alias effect 72
alias frequencies 72
allpass 45
analog to digital conversion 66, 73
analog to digital converter
- dynamic range 77
- saturation 79
- single bit 88
- single slope 74

anti-alias filter 72
attenuation factor Q 9
attenuation spectrum 8

**B**
bilateral z-transform 107

**C**
calibration gain 176, 177, 180
causality 148
Chalfant Valley earthquake 4–8
channel sensitivity 178, 180
circular convolution 116
coding 75
- binary 76
- binary code 75
- decimal 76

convolution 6, 110
convolution theorem
- continuous systems 18
- periodic sequences 116

corner frequency 7
crosstalk 4

**D**
damped harmonic oscillator 51
damping constant 57
dashpot 48, 49
decibel (dB) 35, 78
decimation 87, 117
deconvolution 6, 18, 123, 144
- instability 144
- water level correction 145

delta function 22, 23
- sifting property 23

density 9
derivative 16
difference equation 36
differential equation 37, 38
digital anti-alias filter 94, 117
digitization frequency 67
Dirichlet conditions 97
Discrete Fourier Transform 104, 113
- and z-transform 104

discrete system 95
discretization 66
dynamic range 77

**E**
eigenvalues 19
eigenvector 19
electromagnetic sensor 58
Euler's formulas 54

**F**
Fast Fourier Transform 104
filter
- definition 10
- FIR 153
- FIR filter length 153
- non-recursive 118
- recursive 118
- savings account 36

filter delay 188
FIR filter 118, 119
- maximum phase component 122

- minimum phase component 122
- two-sided impulse response 3
folding frequency 71
format
- GSE 150
- SEED 150
Fourier series 96, 103
Fourier transform 14, 95, 97
- and Laplace transform 15, 98
- and sampling 100
- continuous-time signals 97
- discrete-time signals 99
- inverse 14
- of discrete-time signals 101
- shifting theorem 16
- symmetry properties 98
frequency
- angular 14
- digital 102
frequency band 71
frequency domain 15, 97
frequency response function 5, 19, 25, 37, 45
- and pole position 31
- graphical estimation 31
- graphical estimation, example 33
- graphical estimation, general case 42, 43
- slope 35, 45
functions
- general harmonic 15

G
gain ranging 81
general LTI systems
- phase properties 44
generator constant 175
Gibb's phenomenon 120
GRF 142
ground acceleration 51
ground displacement 51
ground motion amplitude 168
ground velocity 58
group delay 190

H
high frequency decay 9
hypocentral distance 9

I
IASPEI viii
IIR filter 118, 119
impulse response 22
- and region of convergence 28
- left-sided 28
- right-sided 28
- stability 37
impulse response function 24, 25, 27, 37, 40
impulse train 99
inertial force 48
initial value theorem 184
instrument simulation 5
instrumental response 8
integration 16
inverse filter 141
inverse Laplace transform 28
inverse theory 9
inverse z-transform 112
irregularity 147, 148
ISOP ix, x

L
Laplace transform 21
- region of convergence 26
- unilateral 185
least significant bit 76
linear phase 45
linear phase filter 119
linear time invariant system 37
linearity 11
logarithmic decrement 57
lowpass filter 33
- corner frequency 34
LTI system
- general 39

M
magnification curve 170

magnitude
- Wood-Anderson 80
maximum phase 44, 45
maximum phase condition 149
minimum phase 44, 45
minimum phase condition 149
mixed phase 45
modern digital recording systems
- building blocks 88, 93
most significant bit 76

**N**
noise 140
noise amplification 144
normalization factor 178, 180
normalization frequency 178
notch filter 46, 148
nucleation phase 131
Nyquist frequency 72, 85

**O**
onset distortion 191
onset order 184
onset polarity 188
onset time 188
onset time determination
- effect of FIR filters 3
oversampling 81, 84, 87, 117

**P**
Parseval's theorem 84
pendulum seismometer 48
periodic convolution 103
periodogram 85
phase delay 190
PITSA viii
plateau value 7
pole 22
- multiple 45
pole vector 32
poles and zeros 45
polynomial rooting 126
PREPROC xi, 150
programmable gain amplifier 81
pulse duration 195

**Q**
quantization error 77
quantization noise 84
- variance 84
quantum 76

**R**
radiation pattern 9
RC filter 12
- difference equation 36
- differential equation 13
- frequency response function 14
- impulse response 22
- system diagram 13
- transfer function 21
recording system
- types 2
region of convergence 26, 39, 40
- types 41
regularization 147
resolution 77
rise time 195

**S**
sampling 99
sampling frequency 67, 72
sampling process 66
sampling theorem 72
saturation 79
savings account 36
scale factor C (GSE convention) 176
seismic acquisition system
- building blocks 94
seismic moment 9, 201
seismic onset 183
seismic source parameters 3–9
seismometer
- building blocks 48
- damping constant 57
- damping factor 51
- frequency response function 60, 61, 63

- generator constant 62
- mechanical damping 59
- natural period 51
- overdamped 55
- pole position 64
- pole positions 63
- transfer function 63
- underdamped 54

shunt resistance 58
sigma-delta modulation 88
signal front delay 190
signal moment 201
signal restitution 5
signal-to-noise ratio 140
simulation 137
simulation filtering 137
site response 8
source spectrum 8
spectral division 6
spectrum 7
spring 48, 49
stability 139, 149
stabilization
- bilateral 150
- unilateral 149

standard instruments 137
step response function 29
system
- anticausal 28
- block diagram 10
- causal 28
- condition for stability 29
- definition 10
- linear 13
- linear time invariant 14
- stable 44

**T**

temporal aliasing 115
time domain 15, 97
time invariance 11
transfer function 5, 21, 25, 37, 42
- definition 22
- displacement 168
- velocity 168

**U**

unit step function 24

**V**

valid frequency band 141, 146

**W**

water level correction 145
waveform changes 2
Whittaker reconstruction 67, 69, 72
wrap around effect 115

**Z**

z plane 108
- mapping from s-plane 109

zero
- multiple 45

zero padding 116
zero phase 45
zero phase filter 119
z-transform 107
- and Laplace transform 109
- convolution theorem 110
- properties 110, 112
- region of convergence 109
- shifting theorem 110
- time inversions 110
- transfer function 107, 113, 151

MODERN APPROACHES IN GEOPHYSICS

1. E.I. Galperin: *Vertical Seismic Profiling and Its Exploration Potential.* 1985 ISBN 90-277-1450-9
2. E.I. Galperin, I.L. Nersesov and R.M. Galperina: *Borehole Seismology.* 1986 ISBN 90-277-1967-5
3. Jean-Pierre Cordier: *Velocities in Reflection Seismology.* 1985 ISBN 90-277-2024-X
4. Gregg Parkes and Les Hatton: *The Marine Seismic Source.* 1986 ISBN 90-277-2228-5
5. Guust Nolet (ed.): *Seismic Tomography.* 1987 Hb: ISBN 90-277-2521-7; Pb: ISBN 90-277-2583-7
6. N.J. Vlaar, G. Nolet, M.J.R. Wortel and S.A.P.L. Cloetingh (eds.): *Mathematical Geophysics.* 1988 ISBN 90-277-2620-5
7. J. Bonnin, M. Cara, A. Cisternas and R. Fantechi (eds.): *Seismic Hazard in Mediterranean Regions.* 1988 ISBN 90-277-2779-1
8. Paul L. Stoffa (ed.): *Tau-p: A Plane Wave Approach to the Analysis of Seismic Data.* 1989 ISBN 0-7923-0038-6
9. V.I. Keilis-Borok (ed.): *Seismic Surface Waves in a Laterally Inhomogeneous Earth.* 1989 ISBN 0-7923-0044-0
10. V. Babuska and M. Cara: *Seismic Anisotropy in the Earth.* 1991 ISBN 0-7923-1321-6
11. A.I. Shemenda: *Subduction.* Insights from Physical Modeling. 1994 ISBN 0-7923-3042-0
12. O. Diachok, A. Caiti, P. Gerstoft and H. Schmidt (eds.): *Full Field Inversion Methods in Ocean and Seismo-Acoustics.* 1995 ISBN 0-7923-3459-0
13. M. Kelbert and I. Sazonov: *Pulses and Other Wave Processes in Fluids.* An Asymptotical Approach to Initial Problems. 1996 ISBN 0-7923-3928-2
14. B.E. Khesin, V.V. Alexeyev and L.V. Eppelbaum: *Interpretation of Geophysical Fields in Complicated Environments.* 1996 ISBN 0-7923-3964-9
15. F. Scherbaum: *Of Poles and Zeros.* Fundamentals of Digital Seismology. 1996 Hb: ISBN 0-7923-4012-4; Pb: ISBN 0-7923-4013-2
    *2nd Edition.* Hb: ISBN 0-7923-6834-7; Pb: ISBN 0-7923-6835-5
16. J. Koyama: *The Complex Faulting Process of Earthquakes.* 1997 ISBN 0-7923-4499-5
17. L. Tauxe: *Paleomagnetic Principles and Practice.* 1998 ISBN 0-7923-5258-0
18. C.H. Thurber and N. Rabinowitz (eds.): *Advances in Seismic Event Location.* 2000 ISBN 0-7923-6392-2

KLUWER ACADEMIC PUBLISHERS – DORDRECHT / BOSTON / LONDON